Functional Analysis of Syrbactin Synthetase-Like Gene Clusters in *Photorhabdus luminescens* and *Rhizobium sp.* and the Role of Syrbactins in *Pseudomonas syringae* Strains Colonizing Grass Species

Dissertation

zur Erlangung der naturwissenschaftlichen Doktorwürde
(Dr. sc. nat.)

vorgelegt der
Mathematisch-naturwissenschaftlichen Fakultät
der
Universität Zürich

von
Alexey Dudnik
aus
Kasachstan

Promotionskomitee
Prof. Dr. Robert Dudler (Vorsitz und Leitung der Dissertation)
Prof. Dr. Leo Eberl
Prof. Dr. Julia Vorholt

Zürich, 2014

Table of Contents

1. Summary

Syringolin A and glidobactin A are the two well-characterized members of a recently discovered group of proteasome inhibitors called syrbactins. These compounds were first isolated from the plant pathogen *Pseudomonas syringae* pv. *syringae* and the soil bacterium K481-B101 from the order Burkholderiales, respectively. In *P. syringae*, syringolin A acts as a virulence factor and is involved in the suppression of host defense responses. Glidobactin A was first isolated due to its anti-fungal activity, and therefore could potentially play a role in competition with other species. Both compounds were later demonstrated to inhibit the eukaryotic proteasome by irreversibly binding to its catalytic residues. Syrbactins are products of mixed non-ribosomal peptide synthetase/polyketide sythases (NRPS/PKS), which are encoded within gene clusters with characteristic architectures. Extensive analysis of the syringolin synthetase gene cluster and the encoded proteins allowed constructing a complete biosynthetic model for this compound. In the case of glidobactin A, the model is also nearly complete. Homologous glidobactin synthetase-like gene cluster are also found in a number of pathogens, some of which cause lethal diseases. These include the human pathogens *Burkholderia pseudomallei* (the causal agent of melioidosis), *B. oklahomensis* and *Photorhabdus asymbiotica*, as well as the insect pathogen *Ph. luminescens*, suggesting that these pathogens are able to produce glidobactin-like proteasome inhibitors perhaps playing a role in their virulence.

In a first part of this work we determined whether *Ph. luminescens*, the only non-human pathogen in the group, was able to produce a glidobactin-like compound. Compared to K481-B101, the glidobactin synthetase-like gene cluster of *Ph. luminescens* lacks three genes, one of which encodes a putative dioxygenase hypothesized to be involved in the hydroxylation of glidobactin A's macrolactam ring. Unfortunately, like many NRPS-encoding gene clusters, this one turned out to be cryptic and all our attempts to identify conditions suitable for its activation were unsuccessful. Therefore, the gene cluster was cloned and expressed in the heterologous host *P. putida*, upon which production of *bona fide* glidobactin A was detected in culture supernatants. We conclude that the glidobactin synthethase-like gene cluster of *Ph. luminescens* is sufficient for the production of glidobactin A, and that the dioxygenase absent in *Ph. luminescens* may be involved in tail modifications of minor glidobactin variants, which, however, were not detected in our system.

Recently, as described in the second part of the result section, we detected a syringolin synthethase-like gene cluster containing two additional open reading frames of unknown function in *Rhizobium sp.* AP16, a non-pathogenic endophyte isolated from poplar roots. We were able to identify a medium and conditions that induced this strain to produce syringolin A and several minor variants. Interestingly, the relative abundance of several minor syringolin variants was different from that of observed in *P. syringae*, suggesting a difference in amino acid incorporation efficiency by the corresponding NRPS/PKS. Moreover, one of the two adjacent unknown ORFs was co-transcribed with the other syringolin biosynthetic genes and was shown to enhance the overall efficiency of syringolin production. In addition, an unusual duplicated thiolation domain was detected in one the NRPS modules. However, we demonstrated that deleting the second domain does not reduce the yield of syringolin production. In conclusion, AP16 is capable of producing syringolin A. Possible biological roles of this compound in AP16 are discussed.

Because strains producing syringolin do not grow on *Arabidopsis thaliana*, strains able to colonize wheat (*Triticum aestivum*) that naturally contained or did not contain syringolin synthetase gene clusters were chosen to assess contributions of syringolins to virulence in the context of other virulence factors. When tested on wheat, the effect of syringolin production was identified to be dosage dependent, and while production in wild type amounts was found to enhance disease symptom development, overproduction of this compound turned out to be detrimental for the infection process. In order to better understand any potential observations using this model, the genomes of two wheat isolates used in this study were sequenced and annotated. These were *P. syringae* pv. *syringae* (*Psy*) B64, which contains a syringolin synthetase gene cluster, and *Psy* SM, which naturally lacks one. Both genomes contained a relatively small complement of known type III effectors (T3Es), which suggested that other virulence factors might be important. Nevertheless, the type III secretion system was identified to be the key element allowing endophytic growth of these strains. *Psy* SM, which, with seven T3Es, has the smallest T3E repertoire hitherto described, was further characterized by assaying deletions of individual effector genes, or combinations thereof, as well as mutants with disabled exopolysaccharide biosynthetic genes or type VI secretion systems. Surprisingly, despite the small number, the contribution of individual effectors was also relatively small. In addition, the type VI secretion system was demonstrated to be involved in competition with other bacteria, but seemed to play no role in virulence.

Last, the genomes of *Psy* SM and *Psy* B64 were compared with genomes of other *P. syringae* strains isolated from grasses (family Poaceae) in order to find potential adaptations for a lifestyle in monocot hosts. The comparison allowed identifying several conserved genes which might potentially be involved in the protection from host defense responses. Moreover, Poaceae isolates were shown to represent two different phylogenetic lineages, each adopting different strategies for host colonization. The small complement of type III effectors was identified to be a characteristic feature of one of these groups. This analysis provides a foundation for further experiments, which should improve our knowledge of plant-pathogen interactions, in particular for monocot plants.

2. Zusammenfassung

Syringolin A und Glidobactin A gehören zu den am besten charakterisierten Mitgliedern der Syrbactine, einer vor Kurzem entdeckten neuen Strukturklasse von Proteasomeninhibitoren. Diese Verbindungen wurden zuerst aus dem pflanzenpathogenen Bakterium *Pseudomonas syringae* pv. *syringae* und dem Bodenbakterium K481-B101 (Burkholderiales) isoliert. Syringolin A ist ein Virulenzfaktor von *P. syringae* und ist in die Suppression von Abwehrreaktionen der Wirtspflanzen involviert, während Glidobactin A als antifungaler Stoff charakterisiert wurde. Beide Verbindungen inhibieren das eukaryontische Proteasom, indem sie kovalent an das aktive Zentrum aller drei katalytischen Proteasomenuntereinheiten binden. Syrbactine werden von gemischten nicht-ribosomalen Peptidsynthetasen/Polyketidsynthasen (NRPS/PKS) synthetisiert, die von Gengruppen mit charakteristischer Architektur kodiert werden, die es erlaubte, detaillierte Biosynthesemodelle zu postulieren, die zum grossen Teil auch experimentell untermauert sind. Homologe Glidobactin-Synthetasegene wurden auch in den Genomen einer Reihe von pathogenen Bakterien gefunden, wie zum Beispiel in *Burkhoderia pseudomallei*, dem Erreger der der Melioidose, oder im Insektenpathogen *Photorhabdus lumineszens*, was zur Hypothese Anlass gab, dass diese Organismen möglicherweise ebenfalls Proteasomeninhibitoren bilden.

Im ersten Teil dieser Arbeit wude dieser Frage am Beispiel von *Ph. luminescens*, dem einzigen Nichthumanpathogen dieser Gruppe, nachgegangen. Im Vergleich zu K481-B101 fehlen der Gruppe von putativen Glidobactinsynthetasegenen drei Gene, wovon eines aufgrund von Sequeznähnlichkeiten möglicherweise für eine Dioxygenase kodiert, die eventuell für die Einführung der Hydroxylgruppe in die Ringstruktur von Glidobactin A verantworlich ist. Leider erwies sich aber die Gengruppe als kryptisch, d. h. es konnten trotz intensiver Suche keine Bedingungen gefunden werden, unter denen diese Gene transkribiert werden. Deshalb wurde die ganze Gengruppe in ein Cosmid kloniert und in verschiedene *Pseudomonas*-Arten transformiert. Dies führte im Fall von *P. putida* zum Erfolg, und es konnte genuines Glidobactin A aus Zellkulturüberständen isoliert und identifiert werden. Daraus wurde geschlossen, dass die Glidobactinsynthase-ähnliche Gengruppe von K481-B101 zur Glidobactin A-Produktion ausreicht, und dass das nicht vorhandene putative Dioxygenasegen nicht für die Hydroxylierung der Ringstruktur verantwortlich ist, sondern wohl eher in die Modifikation des Fettsäureschwanzes von Glidobactin-Varianten involviert ist.

Wie im zweiten Resultateteil beschrieben wird, haben wir kürzliche eine Syringolinsnythetase-ähnliche Gengruppe mit zwei zusätzlichen Genen unbekannter Funktion in *Rhizobium sp.* AP16, einem nicht pathogenen endophytischen Bakterium, das aus den Wurzeln nordamerikanischer Pappeln isoliert wurde. Es ist gelungen, Kultivationsbedinungen zu finden, unter denen das Bakterium Syringolin A und einige Varianten synthetisiert. Interessanterweise war das Verhältnis zwischen verschiedenen Syringolinvarianten von demjenigen, das in *P. syringae*-Stämmen beobachtet wird, verschieden, was auf unterschiedliche Aminosäurenselektivitäten der entsprechenden NRPS/PKS Module schliessen lässt. Ebenfalls konnte gezeigt werden, dass unter Syringolin-produzierenden Bedingungen auch eines der beiden neuen Gene, welches für eine Typ II Thioesterase kodiert, aktiv transkribiert wurde, und dass dieses Gen die Syntheseeffizienz von Syringolin steigerte. Zusätzlich wurde in einem der NRPS Module eine einzigartige duplizierte Thiolationsdomäne gefunden. Diese sind von PKS und Synthetasen polyunsaturierter Fettsäuren vereinzelt bekannt, wo sie zum Teil die Syntheserate beschleunigen können. Die Analyse von gezielt mutierten Stämmen, in denen die eine oder andere dieser Domänen deletiert war, zeigte aber, dass die Deletion der zweiten Domäne im Vergleich zum Wildtyp keine Veränderung der Syringolinpoduktion verursachte. Zusmmenfassend lässt sich sagen, dass *Rhizobium sp.* AP16 Syringolin A und Nebenvarianten synthetisiert, was darauf hindeutet, dass der Proteasomeinhibitor Syringolin A möglicherweise nicht nur bei Pathogenen, sondern auch bei Endophyten eine Funktion bei der Unterdrückung von Abwehrreaktionden des Wirtes hat.

Bis jetzt konnten leider keine Syringolin-produzierenden Bakterienstämme identifiziert oder konstruiert werden, welche die Modellpflanze *Arabidopsis thaliana* kolonisieren könen. Um die Wirkung von Syringolin A für die Kolonisation einer Wirtspflanze durch ein phytopathogenes Bakterium besser studieren zu können, wurden Stämme ausgewählt, die den Weizen (*Triticum aestivum*) befallen können und die natürlicherweise Syringolin A produzieren konnten oder dazu mangels entsprechender Gene nicht in der Lage waren, wie die Genomsequenzierung und Annotation zweier Stämme, *P. syingae* pv. *syringae* (*Psy*) B64 und *Psy* SM, zeigte. Infektionsexperimente, die im dritte Resultateteil beschrieben werden, zeigten, dass die Wirkung von Syringolinen isofern dosisabhängig war, dass Wiltypmengen zwar die Krankheitsymptome verstärkten, eine Überproduktion aber der Kolonisation der Pflanze sehr abträglich war. Die Genome beider Stämme enthalten erstaunlich wenige bekannte Typ III Effektoren (T3Es), was darauf hindeut, dass auch andere Virulenzfaktoren eine Rolle spielen. Durch gezielte Mutagenes konnte aber gezeigt werden, dass das Typ III Sekretionssystem, das für das Einschleusen der T3Es

in die Wirtszellen verantwortlich ist, auch in diesen Stämmen eine Schlüsselrolle für die Wirtskolonisation einnimmt. *Psy* SM, das mit sieben T3Es das bis heute kleinste bekannte T3E Repertoir aufweist, wurde daher durch Deletion einzelner oder mehrerer T3Es näher untersucht. Zudem wurden Mutanten konstruiert, in denen weitere mögliche Virulenzfaktoren wie die Exopolysaccharidproduktion oder das Typ VI Sekretionssystem defekt waren. Es stellte sich erstaunlicherweise heraus, dass trotz der kleinen Anzahl an T3Es die Deletion einzelner T3E Gene einen relativ kleinen Einfluss hatten. Es zeigte sich auch, dass das T6SS keine messbare Funktion für die Virulenz hatte, wohl aber in die Kompetition mit anderen Bakterien involviert war.

Schliesslich wurden im letzten Resultatekapitel die Genome von *Psy* B64 und *Psy* SM mit allen sieben bekannten Genomen anderer aus Grasarten (Familie Poaceae) isolierter *Pseudomonas* Stämme verglichen, um Hinweise auf Adaptationen für die Kolonisation von Monokotylendonenarten zu erhalten. Es konnten einige konservierte Kandidatengene identifiziert werden, die möglicherweise dem Schutz des Pathogens vor Immunreaktionen monokotyledoner Wirtsarten dienen. Es konnte auch gezeigt werden, dass die aus Gräsern isolierten Stämme verschiedene phylogenetische Linien repräsentierten, welche unterschiedlich Strategien der Wirtskolonisation zu haben scheinen: eine dieser Gruppen zeichnet sich durch eine kleine Anzahl von T3Es aus. Diese Analyse ergibt eine gute Grundlage für weitere Experimente zum Verständnis der Interaktion von phytopathogenen Bakterien mit Wirtspflanzen, insbesondere mit Gräsern.

3. General Introduction

3.1. Plant-microbe interactions

Prokaryotes can be considered as the dominant group of living organisms on Earth. Although unnoticed to the naked eye, when combined together, Bacteria and Archaea contain nearly half of the total carbon found in living organisms (Whitman, 1998). These organisms catalyze unique reactions in the global element cycles, serve as bases in a large number of food chains, take a major part in atmospheric oxygen production, and act as symbionts or pathogens of higher organisms, just to name a few of their important functions. Prokaryotes are able to colonize various habitats and environments, ranging from activated sludge and hydrothermal vents to human gut and vacuoles of sponges (Le Chatelier et al., 2013; Gilbert et al., 2012; Hess, 2004; Jiao and Zheng, 2011; Kämpfer, 1997; Orcutt et al., 2011; Takai and Nakamura, 2011; Williamson, 1979). In this respect, plants are not an exception, and most of the their surfaces are usually colonized by microbes (Philippot et al., 2013; Vorholt, 2012). In addition to epiphytic microbes, i.e. those which live on the surface of a plant, there also exist endophytic bacteria which either colonize intercellular spaces, or form an even tighter interaction by living inside plant cells. These include both symbiotic bacteria, such as nitrogen-fixing root nodule bacteria (Denison and Kiers, 2011; Gyaneshwar et al., 2011), as well as a variety of pathogenic species (Mansfield et al., 2012).

3.2. The plant immune system

Bacteria can enter the plant body either via natural openings, such as stomata, or through wounds. Therefore, one of the challenges to a plant is being able to recognize and fend off any harmful intruder, while at the same time being able to establish a symbiotic relationship with beneficial microorganisms (Pieterse and Dicke, 2007). Like animals, plants also possess complex mechanisms for recognition and containment of pathogen attacks. The plant immune system consists of two major components: the microbe- or pathogen-associated molecular pattern (PAMP)-triggered immunity (PTI) and the effector-triggered immunity (ETI) (Dodds and Rathjen, 2010; Jones and Dangl, 2006).

A major role in PTI is played by a number of receptor kinases functioning as pattern recognition receptors (PRR) located on the plasma membrane, which recognize microbe-derived molecules such

as flagellin, chitin, bacterial elongation factor Tu, and lipopolysaccharides (Boller and Felix, 2009; Dodds and Rathjen, 2010). One of the best studied PRR is FLS2 from *Arabidopsis thaliana*, which recognizes flagellin (Chinchilla et al., 2006). Upon perception of its target ligand, the receptor kinase FLS2 activates a signaling cascade by phosphorylating a cognate mitogen-activated protein kinase (MAP kinase). When the strength of the signaling reaches a certain threshold, it elicits basal defense responses, which include the induction of defense gene expression, production of reactive oxygen species (ROS), detoxifying enzymes and anti-microbial peptides, as well as callose deposition and cell wall re-enforcement (Dodds and Rathjen, 2010; Jones and Dangl, 2006).

In order to be able to survive and reproduce inside its host, plant pathogens have evolved a number of tools for suppressing the basal immune response. The majority of Gram-negative plant pathogens do so using the type III secretion system (T3SS). The T3SS is a protein delivery machinery which uses a structure called injectosome for delivery of effector proteins directly into host cells by puncturing the cell membrane (Büttner, 2012; Galán and Wolf-Watz, 2006). Effector proteins, in turn, interfere with the basal immune response at different stages, thus, preventing PAMP recognition and leading to the so-called effector-triggered susceptibility (ETS) of the host plant.

Throughout their evolution, plants have evolved ways of sensing and recognizing the presence of microbial effectors, which results in ETI, the second level of the plant immune response (Jones and Dangl, 2006). The successful recognition of an effector usually results in localized programmed cell death, also called hypersensitive response (HR). Since the effector composition is variable between different strains and races of microbes, this type of interaction is usually race-specific (Dodds and Rathjen, 2010). The recognition can occur both directly or indirectly and is mediated by the so-called resistance gene (R-gene) products. The majority of know R-genes encode nucleotide-binding domain and leucine-rich repeat (NLR)-containing proteins. With regard to structure, plant NLRs resemble animal NLRs, but are presumed to have an independent evolutionary origin, also because they are absent in unicellular green algae. The number of NLRs is highly variable, ranging from three dozen in papaya (*Carica papaya*) to nearly 500 in rice (*Oryza sativa*) (Jacob et al., 2013). Very often the recognition of microbial effectors occurs directly via binding by their cognate NLRs. Recently, an NLR protein was demonstrated to directly associate with two unrelated fungal effectors (Cesari et al., 2013). Alternatively, recognition can occur indirectly as, for example, seen for the SUMM2 protein in *A. thaliana*, which detects disruption of MEKK1-MKK1/MKK2-MPK4 signaling cascade by microbial effectors (Zhang et al., 2012). It should be noted that NLRs

sometimes recognize small molecules other than proteins, such as for example low-molecular-weight compounds syringolides (Ji et al., 1997; Keen et al., 1990). Microbes, in turn, would either loose the recognized effector or acquire novel effector(s) which would prevent recognition by NLRs. This evolutionary arms race is often referred to as the zig-zag model, where one side gets an advantage in terms of new effectors or R-genes, until the second player would acquire a mean to counteract it (Jones and Dangl, 2006). In conclusion, the combination of PTI and ETI provides rather effective means of perceiving danger and defending the plant from pathogens (Boller and Felix, 2009).

Another important element of plant defense is stomatal immunity. Stomata, being one of the two types of natural openings in plants, are essentially the major gateway for bacterial pathogen entry, which, unlike fungi, are not able to penetrate the plant cell wall directly. Therefore, plants have developed mechanisms for closing the stomata in response to the presence of pathogens (Melotto et al., 2006). The mechanism is essentially a basal defense response, where upon perception of pathogen-born molecules, such as chitosan, by PRRs of the guard cells, a rapid closure of the stomatal aperture is triggered (Lee et al., 1999; Melotto et al., 2008). PAMP-triggered closing of stomata is a complex process which involves signaling by salicylic acid (SA), MAP kinase 3, and absisic acid-mediated signaling (Melotto et al., 2008). Since the closure of stomata greatly limits access to the leaf mesophyll, bacteria have evolved several ways of preventing this (Melotto et al., 2006; Schellenberg et al., 2010).

Another important component of plant defenses is systemic acquired resistance (SAR), which is a combination of induced defense responses that allow long-lasting protection against a broad spectrum of microbes. SAR is usually induced by cell death in distant tissues, which could be caused by both HR and/or pathogen attack. As a result, a number of pathogenesis-related genes (PR genes) are up-regulated in both adjacent and distant tissues (Durrant and Dong, 2004). The spread of SAR is mediated by methyl-SA, as well as by azelaic acid, dihydroabetinal, and glycerol-3-phosphate (Glover et al., 2013). SAR is essentially priming distant tissues in order to prepare them for a potential spread of the attacking pathogen and other secondary infections, which is analogous to interferon-γ signaling between mammalian cells (Conrath et al., 2006). Despite the relatively high energy costs of activation and maintenance of the defense systems, SAR provides an additional level of protection from potential pathogen attacks, in particular during large-scale disease outbreaks.

3.3. Pseudomonas syringae

Strains of the *Pseudomonas syringae* group are causal agents of a variety of plant diseases worldwide. While the first isolate of this species was obtained from a lilac tree (*Syringa vulgaris*) (Palleroni, 2005a), strains of this group have been reported to infect nearly 200 different plant species (Bradbury, 1986). The list includes many agriculturally important crops, such as bean, sugarbeet, tomato, pea, as well as stone fruit, pear, and olive trees. As a consequence, *P. syringae* is one of the best-studied plant pathogens. While some of *P. syringae*-induced diseases are known to recur in the form of outbreaks since a long time (Arnold et al., 2011; Gironde and Manceau, 2012; Martín-Sanz et al., 2013; Ramos et al., 2012), others have only been described recently (Green et al., 2010; McCann et al., 2013). The types of diseases and symptoms caused by *P. syringae* are highly variable and include bleeding canker, bacterial knot, brown spot disease, blast of stone fruit trees, halo blight, apical necrosis, and red streak diseases (Buell et al., 2003; Cazorla et al., 1998; Feil et al., 2005; Green et al., 2010; Joardar et al., 2005; Jones and Benson, 2001; Rahimian, 1995; Ramos et al., 2012; Sinclair et al., 1987).

The taxonomy of *P. syringae* remains under continuous debate. According to the commonly used version, the species is sub-divided into over 50 groups called “pathovars” (Young, 2010; Young et al., 1978). This nomenclature is based on several properties, including host range, disease type, and biochemical characteristics of a strain (Qi et al., 2011; Young, 2010). Several DNA hybridization studies have demonstrated that the species consists of several groups that show a relatively low level of hybridization between one other. However, biochemical and metabolic properties, with a few exceptions, did not allow to elevate these groups to individual species (Gardan et al., 1992, 1999). Currently, there are two commonly used versions of *P. syringae* classification: the first version is based on multi-locus sequence typing (MLST) and divides the species into five phylogenetic groups (Hwang et al., 2005; Sarkar and Guttman, 2004), while the second one is based on DNA hybridization profiles and splits it into nine so-called genomospecies (Gardan et al., 1999).

3.4. Pathogenicity and virulence factors of *P. syringae*

P. syringae is a highly versatile and genetically diverse pathogen, and the observed diversitty is most pronounced at the inter-pathovar level. Furthermore, strains isolated from the same host but different geographic regions might also be quite diverse (McCann et al., 2013; O’Brien et al., 2012; Qi et al., 2011). The highest degree of variation is seen in the complement of virulence factors,

which ultimately shapes the host range, the disease type, and its severity (Lindeberg et al., 2009). In addition to that, some strains possess additional metabolic adaptations that promote their grow in a particular host (Green et al., 2010; McCann et al., 2013; Rodríguez-Palenzuela et al., 2010). For successful survival and reproduction both inside and outside the host, *P. syringae* strains produce various substances that include type III-secreted effector proteins, phytotoxins, exopolymeric substances, and phytohormones (Bender et al., 1999; Denny, 1995; Haapalainen et al., 2012; Lindeberg et al., 2009, 2012; Misas-Villamil et al., 2013; Rodríguez-Moreno et al., 2008; Schellenberg et al., 2010; Yu et al., 1999).

In *P. syringae,* the T3SS is the key factor defining the host range of a strain. It is a true pathogenicity factor, as knocking it out renders the bacterium avirulent in most cases (Lindeberg et al., 2009, 2012). The canonical *hrp*/*hrc*-type T3SS found in the majority of *P. syringae* isolates is encoded by a conserved gene cluster. The gene cluster consists of genes encoding structural proteins, as well as of two loci containing effector-coding genes (Alfano et al., 2000). The location of type III effectors (T3Es) is not limited to these loci, and most of them are actually encoded in other genomic regions (Lindeberg et al., 2008). The majority of T3Es are assumed to be involved in the suppression of plant defense responses (Cunnac et al., 2009; Jones and Dangl, 2006). In addition, some effectors were demonstrated to have a cytotoxic effect (Munkvold et al., 2008; Salomon et al., 2012). Until recently, there were 58 verified effector families (Baltrus et al., 2011), with two new studies further increasing this number to 62 (Matas et al., 2014; Mucyn et al., 2014).

Despite extensive research in this field, the exact mechanism of action for a large portion of T3Es remains known. AvrPtoB is probably the best-characterized effector. It is an E3 ubiqutin ligase, which recognizes and marks for ubiquitin-mediated degradation several basal defense-associated proteins, such as the flagellin recognition receptor FLS2, the chitin receptor CERK1, and the signaling kinase Fen (Gimenez-Ibanez et al., 2009; Göhre et al., 2008; Rosebrock et al., 2007). Another multifunctional effector is the acetyltransferase HopZ1a, which interferes with plant microtubule network formation and vesicle trafficking, as well as with jasmonic acid (JA)-mediated signaling by targeting jasmonate ZIM-domain (JAZ) transcriptional regulators (Jiang et al., 2013; Lee et al., 2012a). JA is involved in several signaling pathways, one of which leads to the suppression of the SA-dependent defense pathway (Derksen et al., 2013). A lot is also known about HopM1, which was demonstrated to destabilize ADP ribosylation factor guanine nucleotide exchange factor (ARF-GEF) proteins, such as AtMIN7, and transcription factor TGA3, therefore

suppressing both PTI and ETI (Gangadharan et al., 2013; Nomura et al., 2006, 2011). The number of type III effector-coding genes is variable among the *P. syringae* strains, and in the majority of strains their number ranges from two to three dozen (Baltrus et al., 2011; McCann et al., 2013; Qi et al., 2011). However, not all of the encoded effectors are essential for full virulence due to functional redundancy (Cunnac et al., 2011). As a result, it is not uncommon that strains isolated to the same host have considerable differences in their effectors sets (Almeida et al., 2009; McCann et al., 2013; O'Brien et al., 2012; Qi et al., 2011). The effector repertoires are under heavy evolutionary pressure (Jones and Dangl, 2006), and thus are being continuously remodeled, which might eventually lead to a change of host specificity. The field of effector biology still attracts a lot of attention, aiming at improving our knowledge of the molecular biology of plant-pathogen interactions.

The *hrp*/*hrc*-type T3SS is not the only one present in *P. syringae* as several strains were recently described to possess a gene cluster encoding a rhizobial-like T3SS (Gazi et al., 2012). Currently, very little is known about this T3SS, and it remains to be discovered whether it is active and plays any role in interactions with the host. In addition, several grass isolates were identified to contain a non-canonical T3SS which is only distantly related to the *hrp*/*hrc*-type T3SS (Clarke et al., 2010). The respective gene cluster is found at a different genomic site and has several conserved genes replaced or missing. From the known T3Es, only homologs of AvrE1 and HopM1 were found in these strains. Moreover, this T3SS appears to have no apparent role during *in planta* growth.

Aside of the T3SS, *P. syringae* strains often produce other substances which enhance virulence and epiphytic fitness. Exopolysaccharides (EPS) are among the best studied virulence factors in *P. syringae*. There are two major EPS detectable in this species: levan and alginate (Osman et al., 1986). Levan is a branched β-polyfructan, which is synthesized from sucrose by a single enzyme called levansucrase (Li and Ullrich, 2001; Osman et al., 1986). There are usually one to three levansucrase genes found *P. syringae* genomes that are located at different genomic sites (Li and Ullrich, 2001). Alginate, in contrast, is a heteropolymer and consists of blocks of non-repeating α-L-glucuronate and β-D-mannuronate, where the latter one gets randomly acetylated (Franklin et al., 2011). Alginate biosynthesis is a multi-step process which involves several consecutively-acting enzymes (Franklin et al., 2011). Most of the biosynthetic genes involved in alginate production are found in a single gene cluster which has an identical architecture in both *P. syringae* and *P. aeruginosa* (Penaloza-Vazquez et al., 1997). The major roles of EPS are assumed to be protection from environmental stress, as well as masking the bacterium from detection by the host.

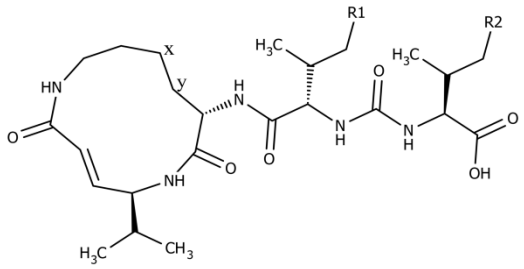
Furthermore, while alginate was demonstrated to enhance epiphytic fitness and osmotolerance (Penaloza-Vazquez et al., 1997; Yu et al., 1999), levan is thought to rather serve as storage polymer (Laue et al., 2006). In addition, genomes of most *P. syringae* strains encode homologs of the *psl* gene cluster from *P. aeruginosa*. Psl is a complex polysaccharide consisting of repeating pentamer units of D-glucose, D-mannose, and L-rhamnose (Byrd et al., 2009) that was shown to be involved in biofilm formation in *P. aeruginosa* (Jackson et al., 2004).

The second important group of virulence factors comprises the phytotoxins. There appears to be no general dependency on phytotoxin production, and while some strains were demonstrated to secrete more than one such compounds, others lack any known phytotoxin biosynthetic genes (Baltrus et al., 2011; Bender et al., 1999). Moreover, strains isolated to the same host sometimes have differences in phytotoxin gene content (Cai et al., 2011; Murillo et al., 2011; Qi et al., 2011). Nevertheless, it is a well-established fact that their presence enhances disease progression and symptom development (Arrebola et al., 2009; Bender et al., 1999; Groll et al., 2008). Strains of *P. syringae* are known to produce several classes of phytotoxins. The grouping is based on targets and mode of action. One group consists of the antimetabolite toxins that inhibit various steps in nitrogen metabolism, which results in chlorosis of surrounding host tissue. It consists of phaseolotoxin, mangotoxin, tabtoxin, and several other uncharacterized toxins. Tabtoxin inhibits glutamine synthetase, mangotoxin targets ornithine acetyltransferase, while phaseolotoxin blocks ornithine carbamoyl transferase (Arrebola et al., 2011; Bender et al., 1999). Another type of phytotoxin is tagetitoxin, which targets chloroplast, bacterial, and some eukaryotic RNA polymerases (Artsimovitch et al., 2011). The third group consists of the lipopeptides syringomycin and syringopeptin, which are assumed to cause necrosis of host tissue by imbedding into the host plasma membrane and inducing ion leakage (Bender et al., 1999). The last group consists of compounds that modulate host defense signaling. One member of this group is the polyketide coronatine, which bears resemblance to the JA (Bender et al., 1999). As mentioned previously, JA counteracts the SA-mediated signaling, which results in down-regulation of host defense responses. Interestingly, coronatine and the T3E HopZ1a share a number of targets, and the coronatine-negative phenotype could be partially rescued by bringing in a copy of HopZ1a (Jiang et al., 2013). Another member is the macrolactam syringolin A, which is a well-characterized inhibitor of the eukaryotic proteasome (Groll et al., 2008). A number of hormone-based signaling pathways in dicot plants require proteasome-mediated protein degradation, including the SA-mediated defense signaling pathway (Kelley and Estelle, 2012). Notably, both coronatine and

syringolin were demonstrated to be involved in counteracting stomatal immunity, thus promoting the entry of bacteria into the host mesophyll (Melotto et al., 2006; Schellenberg et al., 2010).

3.5. Syringolin A

Syringolin A is the product of a mixed non-ribosomal peptide synthetase (NRPS)/polyketide synthetase (PKS) (Amrein et al., 2004; Ramel et al., 2009). The action of the compound was first discovered in a study on acquired resistance in rice (Reimmann et al., 1995). Acquired resistance is known to also be induced upon inoculation with a non-host pathogen, i.e. a pathogen which is able to infect another plant species but not the target plant (Mysore and Ryu, 2004; Senthil-Kumar and Mysore, 2013). In the original work, Reimmann and colleagues were studying *P. syringae*-induced acquired resistance to *Pyricularia oryzae* in rice (Smith and Métraux, 1991). Among up-regulated PR-gene transcripts, there was one which was detected only upon infiltration with a certain group of *P. syringae* strains. The gene was named *pir7b*, and was subsequently determined to encode an α/β-fold hydrolase (Wäspi et al., 1998a). The induction of *pir7b* was also observed with cell-free culture supernatants, which suggested that its elicitor was a secreted molecule. The compound was further isolated and its chemical structure was resolved (Wäspi et al., 1998b). Subsequent work aiming at discovering the mode of action of syringolin A revealed that the compound irreversibly inhibits all three proteolytic activities (i.e. the caspase-, trypsin-, and chymotrypsin-like activity) of the eukaryotic proteasome by covalent bond formation with the corresponding catalytic residues (Groll et al., 2008). In addition, the compound was demonstrated to promote disease symptom development on bean (*Phaseolus vulgaris*) and to enhance wound entry and lesion spreading on *Nicotiana benthamiana* (Groll et al., 2008; Misas-Villamil et al., 2013).



Syringolin	R1	R2	x-y
A	H	H	-CH=CH-
B	H	H	-CH_2-CH_2-
C	CH_3	H	-CH=CH-
D	H	CH_3	-CH=CH-
E	H	CH_3	-CH_2-CH_2-

Figure 1. Chemical structures of syringolin A and its minor variants

In addition to syringolin A, several other minor variants were identified in culture supernatants of *P. syringae* (Wäspi et al., 1999). All identified compounds contain a twelve-membered macrolactam ring made out of two non-proteinogenic aminoacids: 5-methyl-4-amino-2-hexenoic acid and 3,4-dehydrolysine (Figure 1). Syringolin B and E have the latter residue substituted with lysine. The macrolactam ring is connected to a tail, which in the case of syrigolin A consists of two valine residues which are connected with each other via an unusual ureido linkage. Other minor syringolin variants have one or both valine residues replaced with isoleucine.

Genes that are required for syringolin A biosynthesis are organized in a gene cluster, which has been cloned and characterized (Amrein et al., 2004; Ramel et al., 2009). Among the sequenced *P. syringae* strains, the gene cluster is always found at the same chromosomal location and consists of five open reading frames (ORFs), *sylA* through *sylE*. The product of the *sylA* gene is a LuxR family helix-turn-helix DNA-binding protein which acts as a transcriptional regulator of the gene cluster (Ramel et al., 2012). The *sylB* gene encodes a putative fatty acid desaturase which is assumed mediate the conversion of lysine to 3,4-dehydrolysine in the ring structure. The product of the *sylC* gene is a NRPS enzyme, while the *sylD* gene encodes a mixed NRPS/PKS, both of which play the major role in syringolin biosynthesis. The *sylE* gene encodes a putative syringolin efflux transporter (Amrein et al., 2004).

NRPS are large, often multi-modular, enzymes found exclusively in bacteria and fungi. These enzymes catalyze the biosynthesis of peptides and their derivatives based on the thiotemplate mechanism, which does not require an mRNA template or any other component of the standard protein synthesis machinery (Stachelhaus and Marahiel, 1995). NRPS-based biosynthesis allows greater flexibility in terms of substrate and end-product modifications, thus going far beyond the 20 proteinogenic amino acids. These enzymes are involved in production of a large number compounds, such as antibiotics, siderophores, pigments, anti-cancer agents, toxins, and immunosuppressants (Newman et al., 2003; Nikolouli and Mossialos, 2012; Strieker et al., 2010). Apart from syringolin A, several other secondary metabolites of *P. syringae* are produced by NRPS enzymes, such as the phytotoxins syringomycin and syringopeptin, the surfactant syringafactin, and the siderophore pyoverdine (Bender et al., 1999; Burch et al., 2012; Owen and Ackerley, 2011). There are also number NRPS-coding genes of unknown function present in the genomes of different strains. Such cryptic genes and gene clusters are found in many other bacterial genomes, where up to now it was not possible to identify suitable conditions to induce their expression (Challis, 2008).

A typical NRPS consists of an initiation and one or more elongation modules. If several NRPS subunits act consequently one after another, only the first in sequence would contain the initiation module. Elongation modules usually contain three domains: an adenylation domain (A-domain), which recognizes and activates a target amino acid, a condensation domain (C-domain) which catalyzes peptide bond formation, and a peptidyl carrier protein (PCP) domain, which holds the activated amino acid and transfers it to the next module. The initiation module is there to begin the synthesis by recruiting the first amino acid and, consequently, usually lacks a C-domain. There are often accessory enzymes participating in this process which are responsible for maturation of the product. NRPS-based biosynthesis is terminated at a thioestherase (TE) domain located adjacent to the last PCP domain, which facilitates release of the newly synthesized peptide from the NRPS (Finking and Marahiel, 2004).

Type I PKS are as well modular enzymes which catalyze a reaction similar to fatty acid chain elongation. There are at least three domains present in each module: acyl transferase (AT), ketosynthetase (KS), and acyl carrier (ACP) domains. While the ACP domain holds the newly synthesized molecule, the AT domain incorporates malonyl or methylmalonyl-CoA, and the KS domain catalyzes C-C bond formation (Nikolouli and Mossialos, 2012). PKS modules can contain additional domains, such as enoyl-ACP recutase (ER), ketoacyl-ACP reductase (KR), or β-hydroxyacyl-ACP dehydratase (DH) domains (Hopwood, 1997).

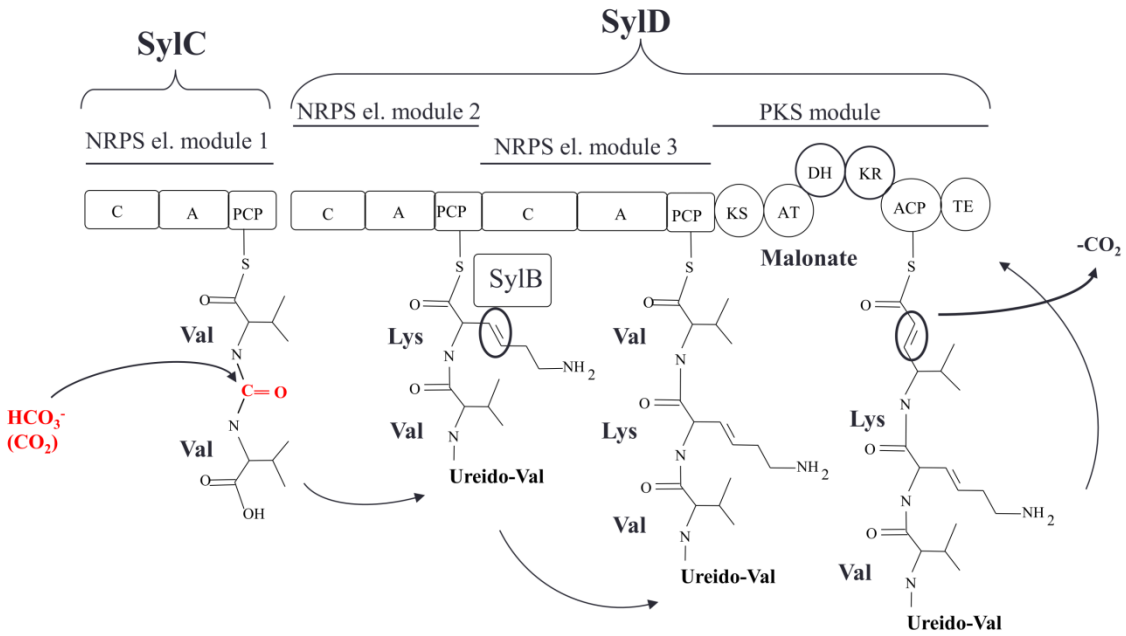


Figure 2. Model for biosynthesis of syringolin A. Modified from Dudler (2013).

Based on both the sequence and the architecture of the *syl* gene cluster, an experimentally-supported biosynthesis model of syringolin A was proposed which explains all structural features of the molecule (Amrein et al., 2004; Imker et al., 2009; Ramel et al., 2009; Wuest et al., 2011). The model is outlined in Figure 2. The minor syringolin variants are the result of incomplete lysine desaturation by SylB and a relaxed specificity of the SylC NRPS module.

3.6. Syrbactins

Glidobactin A and its minor variants are a group of compounds that is structurally-related to syringolins. Glidobactins were originally isolated from soil, and later on were attributed to strain K481-B101 from the order Burkholderiales (Oka et al., 1988a, 1988b). Together, syringolins and glidobactins form a novel class of proteasome inhibitors termed syrbactins (Schellenberg et al., 2007). The macrolactam ring of glidobactins consists of 4(S)-amino-2(E)-pentenoic acid and erythro-4-hydroxy-L-lysine, which is attached to threonine acylated with a fatty acid tail (Figure 3). The other glidobactin variants differ mostly by the length and saturation of the fatty acid part (Oka et al., 1988b, 1988c). Glidobactin A was isolated due to its anti-fungal properties, and later on was also demonstrated to be a proteasome inhibitor (Groll et al., 2008). The respective biosynthetic gene cluster has as well been sequenced and characterized (Figure 4) (Schellenberg et al., 2007).

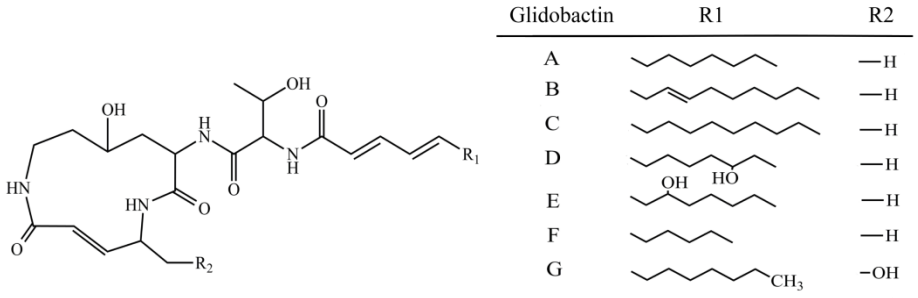


Figure 3. Chemical structures of glidobactin A and its minor variants. Modified from Dudnik, et al (2013).

Interestingly, homologous gene clusters with a nearly-identical architecture were found in several other unrelated bacterial species, including *Burkholderia oklahomensis*, *B. pseudomallei* (both β-proteobacteria), *Photorhabdus luminescens* and *Ph. asymbiotica* (both γ-proteobacteria), which were therefore hypothesized to be capable of producing glidobactin-like compounds (Schellenberg et al., 2007). Interestingly, with the exception of K481-B101, all other bacteria which possess the glidobactin synthatase gene cluster are specialized pathogens: *B. pseudomallei* is the causal agent of melioidosis in humans, *Ph. asymbiotica* and *B. oklahomensis* are associated with human wound

infections, and *Ph. luminescens* is a deadly insect pathogen (Glass et al., 2006; Joyce et al., 2006). In addition, a homologous but distinct gene cluster is also found in a causal agent of glanders in horses, mules, and humans, *B. mallei*. However, in all of the currently sequenced *B. mallei* isolates the gene cluster contains a transposon insertion which disrupts one of the NRPS-coding genes (Figure 4). Therefore it is possible that these pathogens produce syrbactin-like proteasome inhibitors which may play a role in their virulence. More recently, we identified homologs of *sylB-sylE* genes organized identically as in *P. syringae* in the the genome of the α-proteobacterium *Rhizobium sp.* AP16. This bacterium, in contrast, is not a pathogen, but an endophyte isolated from the roots of eastern cottonwood (*Populus deltoides*).

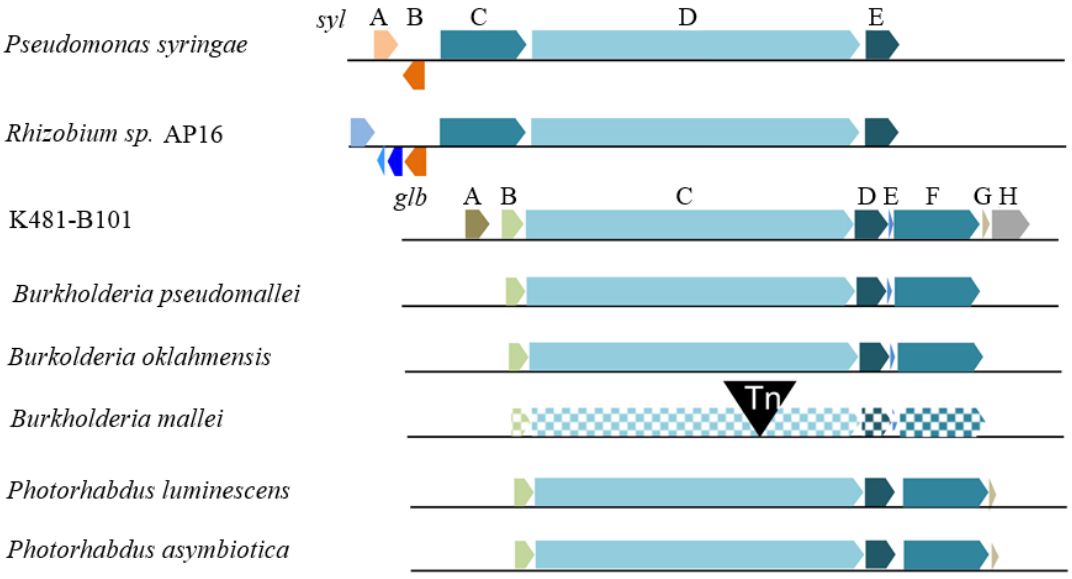


Figure 4. Syringolin synthethase-like gene clusters identified among currently sequenced bacterial genomes. Modified from Dudler (2013).

3.7. Aims of the thesis

This work encompassed several projects which were aiming at improving our knowledge of the biology of syrbactins. The results obtained constitute the contents of Chapters 4, 5, 6, 7, as well as of Appendixes 1 and 2.

The work presented in chapter 4 was aiming at demonstrating whether pathogens with syrbactin synthetase gene cluster are actually able to produce a glidobactin-like compound. Due to fact that *B. pseudomallei*, *B. oklahomensis*, and *Ph. asymbiotica* are human pathogens and therefore require special handling, the work has been carried out with *Ph. luminescens*. In addition, since the

respective gene cluster from *Ph. luminescens* is not identical to that of K481-B101 and several accessory genes are missing, we were hoping to elucidate their function based on potential differences in chemical structure of glidobactin-like compounds produced by the two organisms, thereby improving our knowledge of structure-function relationships of individual components of syrbactin synthetase gene clusters.

The aim of chapter 5 was the characterization of the syringolin synthetase-like gene cluster from the root endophyte *Rhizobium sp.* AP16 (PMI03_05099-PMI03_05102) and its potential role in host colonization. This species is of a particular interest since it is one of the two non-pathogens known to contain a syrbactin synthetase gene cluster. However, during the early stages of colonization, endophytes also encounter plant defense responses, but suppress them with the help of nodulation factors (Shaw and Long, 2003). The genome of AP16 lacks the conventional *nod* genes, and therefore this bacterium might be using syringolin for this purpose. Moreover, the gene cluster contains two additional genes, PMI03_05097 and PMI03_05098, located immediately downstream of the *sylB* homolog (PMI03_05099), which encode a putative tautomerase and a putative type II NRPS-associated thioestherase. In addition, the product of PMI03_05100 (*sylC* homolog) contains tandem thiolation (PCP) domains, a feature normally not found in NRPS that is sometimes seen in PKS, where it was demonstrated to enhance the biosynthesis efficiency (Gu et al., 2011). Therefore, implications of these additional features were investigated.

In chapter 6 we were aiming at the identification of key virulence determinants in *P. syringae* strains which are characterized by a relatively small complement of T3Es. The project has begun as a part of an effort to establish an infection model for syringolin in wheat. Our attempts to obtain a syringolin-producing derivative of a strain pathogenic to *Arabidopsis thaliana* were unsuccessful, and thus, we decided to switch to another common research target, for which we had strains which naturally produce syringolin, as well as a strain which naturally lacks the phytotoxin. Thus, the influence of syringolin production on mesophyll colonization and symptom development was assayed using different inoculation techniques. At the same time, the genomes of the selected strains were sequenced and analyzed. It turned out that both genomes encode a relatively small number of known T3Es, which may suggest that virulence factors other than T3Es are important. Therefore, the first step undertaken was to investigate the importance of the T3SS and of the global virulence regulating two-component system GacA/GacS for colonization of the wheat leaf mesophyll by these

bacteria. Next, we investigated individual contributions of several selected T3Es, as well as of exopolysaccharides and the type VI secretion system.

The aim of chapter 7 was to identify host-specific adaptations of *P. syringae* strains isolated from grasses (family Poaceae). The majority of research projects on this organism involve strains that were isolated from dicot plants, thus leaving the field of monocot-*P. syringae* interactions largely unexplored. We decided to perform a global comparison using genome sequences of *P. syringae* strains isolated from wheat, rice, barley, and proso millet, aiming at the identification of conserved genes found uniquely in these isolates. Products of such genes are likely to be a part of the adaptation of these bacteria to a lifestyle associated with Poaceae family members. In addition, type III effector repertoires, phytotoxin and mobile genetic element compositions, type VI secretion-related gene content, and other virulence-associated traits were analyzed and compared.

Appendixes 1 and 2 contain genome announcements describing properties of the two *P. syringae* genomes sequenced by us: *P. syringae* pv. *syringae* SM and *P. syringae* pv. *syringae* B64.

4. Heterologous Expression of a *Photorhabdus luminescens* Syrbactin-Like Gene Cluster Results in Production of the Potent Proteasome Inhibitor Glidobactin A

Alexey Dudnik[1], Laurent Bigler[2], and Robert Dudler[1,3]

Microbiological Research

168(2):73-76

doi: 10.1016/j.micres.2012.09.006

Published: 22 February 2013

[1] Institute of Plant Biology, University of Zurich, Zollikerstrasse 107, CH-8008 Zurich, Switzerland

[2] Institute of Organic Chemistry, University of Zurich, Winterthurerstrasse 190, CH-8057 Zurich, Switzerland

[3] Corresponding author

4.1. Abstract

Syrbactins are cyclic peptide derivatives which are known to inhibit the eukaryotic proteasome by irreversible covalent binding to its catalytic sites. The only two members of this family characterized to date, syringolin A and glidobactin A, are secreted by certain strains of *Pseudomonas syringae* pv. *syringae* and strain K481-B101 from the order *Burkholderiales*, respectively. Syrbactins are the products of mixed non-ribosomal peptide/polyketide synthases encoded by gene clusters with a characteristic architecture. Similar, but not identical gene clusters are present in several other bacterial genomes, including that of *Photorhabdus luminescens* subsp. *laumondii* TT01, which is therefore hypothesized to be able to produce a syrbactin-type proteasome inhibitor. Here we report the cloning of the putative syrbactins synthetase encoding gene cluster of *Ph. luminescens* into a cosmid vector and its heterologous expression in *Pseudomonas putida.* Analysis of culture supernatants of transformed *P. putida* by HPLC and mass spectrometry revealed the presence of glidobactin A, indicating that the syrbactins-like gene cluster of *Ph. luminescens* encodes a glidobactin A synthetase and that this organism has the capacity to synthesize glidobactin A.

4.2. Introduction

Syrbactins are a structural class of potent proteasome inhibitors that covalently and irreversibly bind to the active site threonine residues of catalytic proteasome subunits (Groll et al., 2008; Krahn et al., 2011). Syringolin A (Figure 1), one of the two well characterized syrbactins to date, is secreted by strains of the phytopathogenic bacterium *P. syringae* pv. *syringae* (Wäspi et al., 1998b). It is the major variant of a small family of closely related compounds thought to be the products of the same synthetase (Wäspi et al., 1999), which is a mixed non-ribosomal peptide synthetase/polyketide synthetase (NRPS/PKS) encoded by the *sylA-sylE* gene cluster (Amrein et al., 2004). Analysis of the gene cluster, in particular of the architecture of the *sylC* and *sylD* genes, have allowed to postulate a biosynthesis model which completely explains the tripeptide part containing the ring structure (Amrein et al., 2004). Recently, other experiments also shed light on the biosynthesis of the tail part of syringolin A, including its unusual ureido group (Imker et al., 2009; Ramel et al., 2009; Wuest et al., 2011). Glidobactin A, the other well characterized syrbactin member, has a twelve-membered ringstructure similar to the one of syringolin A, but exhibits an acylated tail (Figure 1A). It is secreted by an unknown species belonging to the order *Burkholderiales* and is the product of a

synthetase encoded by the *glbA-glbH* gene cluster, which is homologous to the syringolin gene cluster (Figure 1B) (Schellenberg et al., 2007). Like in the case of syringolin A, the architecture of the *sylC* and *sylD* homologs *glbF* and *glbC*, suggested a glidobactin A biosynthesis model which completely explains the twelve-membered ring structure with the exception of the hydroxyl group (Schellenberg et al., 2007). Recently, the acylation of the N-terminal threonine residue has been shown to be mediated by the *glbF* gene product, which, in addition to activating the threonine residue, also mediates its acylation using acyl-CoA as a donor (Imker et al., 2010). Glidobactin A is the major variant of a family of compounds also named cepafungins that differ mostly in the structure of the fatty acid tail (Oka et al., 1988a, 1988c; Terui et al., 1990).

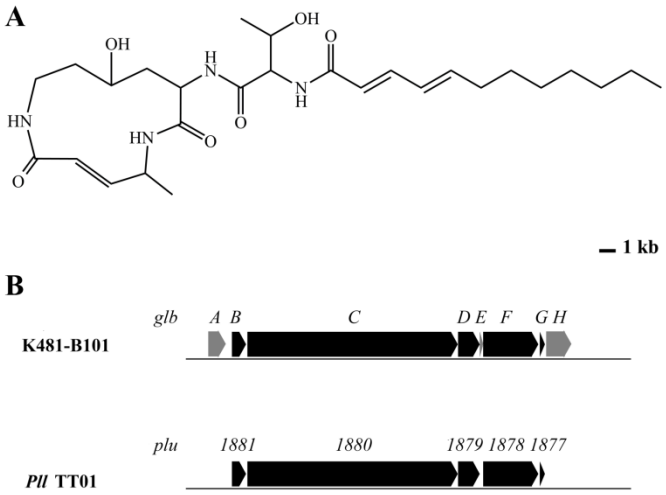


Figure 1. (A) Structural formula of glidobactin A. The compound is a tripeptide derivative with a 12-membered macrolactam ring and a fatty acid tail attached to the N-terminal threonine. (B) Alignment of the glidobactin synthetase gene cluster from K481-B101 (*glbA-H*) with the homologous gene cluster in *Ph. luminescens* ssp. *laumondii* TT01 (*plu1881-1877*). Vectored boxes indicate open reading frames (ORFs). Homologous genes are colored in black, whereas those that are absent in *Pll* TT01 are given in grey.

Among the bacteria whose genomes have been completely sequenced, sequence comparison revealed gene clusters similar to *glbA-glbH* in the insect pathogen *Photorhabdus luminescens* as well as in the human pathogens *Photorhabdus asymbiotica*, *Burkholderia pseudomallei* (causal agent of melioidosis), and *B. oklahomensis*, while *B. mallei* strains contain a disrupted cluster (Schellenberg et al., 2007). In comparison to the *glb* gene cluster, these organisms lack *glbA* and *glbH* homologs, and in some cases also a *glbE* and a *glbG* homologue (Figure 1B). While *glbA* encodes a LysR-type transcription factor and *glbE* an MbtH-like peptide, both of which are

irrelevant for structural considerations (Lautru et al., 2007; Rackham et al., 2010; Schellenberg et al., 2007; Zhang et al., 2010b), the functions of *glbG* and *glbH* are hitherto unknown.

We were interested to compare the structure of the *Ph. luminescens* syrbactins-like gene cluster, which encompasses genes *plu1881-1877*, and the structure of the product of the synthetase encoded by these genes in order to improve our ability to derive product structure from the genes and their architecture. *Plu1881-1877*are homologous to *glbB-glbG*, except that the MbtH-like protein encoded by *glbE* is missing (Figure 1B). However, all our efforts to isolate putative syrbactins from *Ph. luminescens* spp. *laumondii* (*Pll*) TT01 grown under a variety of conditions failed. Thus, we decided to clone *plu1881-1877* and express the gene cluster in a heterologous system. Here we show that *P. putida* harboring a cosmid-borne *plu1881-1877* gene cluster secretes *bona fide* glidobactin A, suggesting the *Pll* is capable to synthesize this potent proteasome inhibitor.

4.3. Materials and Methods

Bacterial strains and culture conditions. Unless stated otherwise, *Escherichia coli* strains, *P. putida* P3 (Senior et al., 1976), and *Ph. luminescens* spp. *laumondii* TT01 (accession no. DSM 15139, German Collection of Microorganisms and Cell Cultures (DSMZ,)) were grown in LB medium on a shaker at 220 rpm at 37°C, 28°C, or 30°C, respectively.

Cloning of the *plu1881-1877* gene cluster and transfer to *P. putida* P3. If not stated otherwise, standard procedures were used (Ausubel et al., 1998). Genomic DNA of *P. luminescence* subsp. *laumondii* TT01 was isolated using Genomic tip-100/G columns (Qiagen, Hilden, Germany) according to instructions of the manufacturer. Ten µg of genomic DNA was digested with the locally unique restriction enzymes SalI and XbaI (positions 2242496 and 2216847, respectively; accession no. NC_005126), separated on a 0.5% agarose gel at 4° C, blotted onto a nylon membrane, and hybridized to a ^{32}P-labeled probe amplified from *plu1880* with primers plu1880_f and plu1880_r (Table 2) in order to verify the expected fragment size. DNA fragments larger than 20 kb were cut out of an identical gel, purified as described (Ramel et al., 2009), and cloned into HindIII/BamHI digested broad-host range cosmid pLAFR3 (Staskawicz et al., 1987) with the help of adapters obtained by annealing oligonucleotides Hind-Sal_Adapt_H/ Hind-Sal-Adapt-S and Xba-Bam-Adapt-X/Xba-Bam-Adapt-B, respectively (Table 1). After the adapters were ligated to the digested vector at 4° C overnight, the modified vector was separated from oligonucleotides by agarose gel electrophoresis and purified using Wizard® *SV* Gel and PCR Clean-Up System

(Promega, Madison, WI, USA). Isolated genomic DNA fragments (>20 kb) were then ligated with the modified vector for two days at 4° C. The products were concentrated by ethanol precipitation and packaged into λ-phage particles using Gigapack III Plus Packaging Extract (Agilent Technologies, Basel, Switzerland). The library was plated on *E. coli* XL-1 Blue MRA (Stratagene, La Jolla, California) and screened according to instructions of the manufacturer using the ^{32}P-labeled *plu1880* probe described above. Positive clones were isolated and completeness of the *plu1881-1877* gene cluster was verified by PCR using primers amplifying the left and right end of the cluster (primers plu1881_f/plu1881_r and plu1877_f/plu1877_r, respectively; Table 1). The recombinant cosmid was named pL3plu and transferred from *E. coli* XL-1 Blue MRA to *P. putida* P3 by tri-parental mating using the *E. coli* helper strain HB101 (pRK600) (Christensen et al., 1999).

Table 1. List of oligonucleotides used

Name	Sequence	Overhang
Hind-Sal-Adapt-H	5'-AGCTTGGAAAGCTT	HindIII
Hind-Sal-Adapt-S	5'-TCGACAAGCTTTCCA	SalI
Xba-Bam-Adapt-X	5'-CTAGACATGTCAGGAG	XbaI
Xba-Bam-Adapt-B	5'-GATCCTCCTGACATGT	BamHI
plu1881_ f	5'-TATGCGTTGTACCACCTTCG	-
plu1881_r	5'-GCCGTTAGATCGATTTGGTG	-
plu1880_f	5'-ATAAACTCCAGTGCGCCATC	-
plu1880_r	5'-ATGTGGATTGTCCCTTCTGC	-
plu1877_f	5'-TTCAAATTTAAAGTAATCGCTGAGG	-
plu1877_r	5'-CGCAGAACTTCTTAGCTCAATG	-

Preparation and analysis of methanolic extracts. Five mL SRM_{AF} medium (Mo and Gross, 1991) containing 10 µg/mL tetracycline were inoculated with *P. putida* and shaken at 220 rpm over night at 28°C. The culture was used to inoculate 100 mL SRM_{AF} medium. These cultures were grown at 28° C either for 2 days at 220 rpm (shaken cultures), or without agitation for 5 days (still cultures). After centrifugation at 10 000 x g for 10 min, pelleted cells were re-suspended in 90% HPLC-grade methanol and shaken at 250 rpm for 2 hours at room temperature. After centrifugation at 10 000 x g for 15 min, supernatants were filtered through 0.22 µm Steritop filters (Millipore, Molsheim, France) and stored at +4°C until used.

HPLC analysis was performed as described previously (Titus and Roundy, 1990) with some modifications. Extracts were concentrated 10-fold by complete evaporation and dissolution in 50% methanol followed by filtration through Millex GP filter units (0.22 µm pore size; Millipore, Molsheim, France). Twenty µL of filtrate was injected into a 250 mm x 4.6 mm Hypersil ODS 5 µm column (Dr. Maisch GmbH, Ammerbuch, Germany) which was connected to UltiMate 3000 HPLC system (Thermo Fisher, Olten, Switzerland). Equilibration of the column and isocratic separation of components was performed with a methanol:water (3:1) solution at a flow rate of 1 mL/min.

Methanolic extracts (50% methanol) were diluted 10:1 or 100:1 with CH_3CN and analyzed with a Waters Acquity UPLC system (Waters, Milford, USA) connected to a Bruker maXis QTof high-resolution mass spectrometer (Bruker Daltonics, Bremen, Germany). An Acquity BEH C18 HPLC column (1.7 µm, 2.1x100 mm fitted with a 2x2 mm guard column) has been used with a mixture of H_2O + 0.1% HCOOH (A) and CH_3CN + 0.1% HCOOH (B) solvent (0.3 ml flow rate, linear gradient from 70 to 100% B within 3 min followed by flushing with 100% B for 1 min). The mass spectrometer was operated in the positive electrospray ionization mode at 4,000 V capillary voltage, -500 V endplate offset, with a N_2 nebulizer pressure of 1.6 bar and dry gas flow of 8 l min^{-1} at 200°C. MS acquisitions were performed in the mass range from *m/z* 50 to 1,200 at 20,000 resolution (full width at half maximum) and 1.5 scan s^{-1}. Masses were calibrated below 2 ppm accuracy with a 2 mM solution of sodium formate over *m/z* 158 up to 1450 mass range prior to analysis. The presence of glidobactin A was confirmed by calculation of extracted ion chromatograms of the $[M + H]^+$ signal at *m/z* 521.3333±0.05.

4.4. Results

In order to identify compounds produced by the *plu1881-1877*-encoded enzymes, a heterologous expression system was explored using *P. putida* strain P3, which has been successfully employed previously for the expression of the syringolin synthase cluster (*syl*) (Ramel et al., 2009). Analysis of the *Pll* TT01 genomic region containing *plu1881-1877* revealed unique SalI and Xba I restriction sites, with the SalI site being 2.54 kb upstream of *plu1881* and the XbaI located 4.77 kb downstream of *plu1877*. These restriction sites, which delimit a 25.65 kb fragment, completely contain *plu1881-1877*. The presence and uniqueness of these sites was confirmed by DNA gel blot analysis of SalI/XbaI-digested genomic DNA of *Pll* TT01 using a *plu1880*-derived hybridization probe (data not shown). A sub-genomic library was then prepared from SalI/XbaI-digested genomic DNA which was size-fractionated by separation on an agarose gel. Fragments larger than 20 kb were eluted and

cloned into the broad-host range cosmid vector pLAFR3, packaged into λ-phage particles, and transfected into *E. coli*. The library was screened by hybridization with a *plu1880*-derived probe.

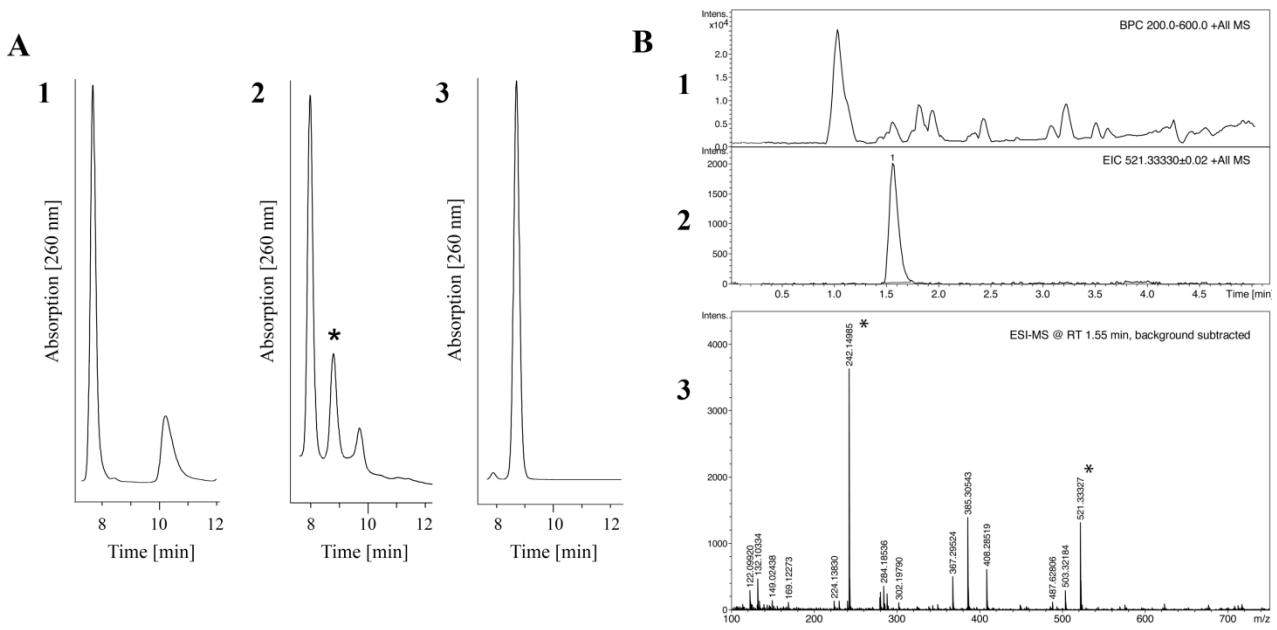


Figure 2. (**A**) HPLC profiles of wild-type *P. putida* extract (1), extract from *P. putida* expressing the *plu1881-1877* gene cluster (2), and isolated glidobactin A standard (3). (**B**) Total ion UPLC chromatogram of methanolic extract of *P. putida* carrying the *plu1881-1877* gene cluster (a). Magnified peak with retention time (RT) of 1.55 min that is expected to contain glidobactin (2). UPLC–MS spectrum at RT 1.55 min (3). The two signals marked with asterisks correspond to protonated glidobactin ($C_{27}H_{45}N_4O_6$, m/z_{exp} 521.33327, m/z_{calc}521.33336, 0.18 ppm error) and a glidobactin fragment generated in the ion source ($C_{11}H_{20}N_3O_3$, m/z_{calc} 242.14992, m/z_{exp}242.14986, 0.24 ppm). The latter corresponds to the macrolactam ring of glidobactin without the extracyclic threonine and fatty acid tail.

Positive colonies were isolated and presence of the entire cluster was the verified by PCR using primers amplifying the 5'-end of *plu1881* and 3'-end of *plu1877*. The recombinant cosmids of positive clones were transferred into *P. putida* P3. Presence of the cluster in transconjugant candidates was verified by PCR using *plu1881*, *plu1880* and *plu1877*-derived primers. Metabolite production was then tested by growing bacteria in a minimal medium either under constant shaking for two days, or as still cultures for five days. Methanolic extracts of pelleted cells were prepared and analyzed by HPLC. Comparison of HPLC profiles from transformed bacteria carrying the *plu1881-1877* gene cluster with profiles from the untransformed control strain reproducibly revealed a differential peak when cultures were grown under still culture conditions. As shown in the Figure 2A, conditioned media of the transformed *P. putida* strain revealed a peak with a retention time identical to the one of isolated glidobactin A (8.7 min). LC-HR-ESI-MS of extracts revealed a quasi-

molecular ion with an *m/z* of 521.33327, corresponding to a formula of $C_{27}H_{45}N_4O_6$ (hydrogen adduct, *m/z* 521.33336, SD of 0.18 ppm), which is identical to the one of the hydrogen adduct of glidobactin A (Figure 2B). Thus, we conclude that the product of the NRPS/PKS encoded by the *plu1881-1877* gene cluster of *Pll* TT01 heterologously expressed in *P. putida* is glidobactin A, and that *Pll TT01* is capable to produce this potent proteasome inhibitor.

4.5. Discussion

We were hitherto unable to detect glidobactin A or similar compounds in *Pll* cultures grown in various media and conditions either by HPLC or by proteasome inhibition assays (unpublished results). Conditions explored included still and shaken cultures in minimal and insect culture media supplemented with insect hemolymph or proline, ingredients which have been shown to activate expression of virulence-associated genes in *Pll* TT01(Crawford et al., 2010). While these compounds were indeed able to slightly induce expression of the cluster, the level of production was apparently still below detection limits (data not shown). It is a notorious problem that for many NRPS and PKS gene clusters identified by genome sequencing, the products synthesized by the encoded enzymes remain unknown because these genes are not expressed under the culture conditions used or explored. In contrast to the glidobactin A gene cluster of K481-B101, the *Pll* syrbactin gene cluster does not encompass a transcriptional activator gene, suggesting that a regulator encoded elsewhere in the genome is involved. As to why the *Pll* gene cluster is active in *P. putida*, we can only speculate. One possibility is that the *Pll* gene cluster is activated in this organism by a regulator active under the culture conditions employed. *Pll* may also harbor a repressor of the syrbactin gene cluster that is absent in *P. putida*. In addition, the fact that multiple copies of the cosmid carrying the *Pll* gene cluster likely are present may also play a role.

Comparison of the amino acid residues forming the binding pocket of the NRPS modules encoded by *plu1880* and *plu1878* with the ones of their homologues of the glidobactin A producing strain K481-B101 suggests that the former activate the same amino acids as the modules present in GlbC and GlbF, respectively (Table 2). However, the detection of *bona fide* glidobactin A in culture supernatants of *P. putida* carrying the *plu1881-1877* gene cluster was unexpected because the gene cluster lacks a *glbH* homologue. The product of this gene is hypothesized to play a role in features of glidobactin A left unexplained by the biosynthesis model (Imker et al., 2010; Schellenberg et al., 2007), *i.e.* in the hydroxylation of the lysine residue, or in the modifications exhibited by the minor

Table 2. Prediction of amino acid specificity of NRPS modules encoded by *glbC/glbF* and *plu1880/plu1878*

Module	Position[a]										Observed	Predicted[b]
	235	236	239	278	299	301	322	330	331	517		
GlbF	D	F	W	N	I	G	M	V	H	K	Thr	Thr
Plu1878	D	F	W	N	I	G	M	V	H	K	Thr	Thr
GlbC-1	D	L	G	D	V	G	S	I	D	K	Lys	Lys
Plu1880-1	D	L	G	D	V	G	S	I	D	K	Lys	Lys
GlbC-2	D	V	G	W	I	A	G	I	V	K	Ala	-
Plu1880-2	D	V	G	W	I	T	G	I	V	K	Ala	-

[a] Position of amino acid residues (given one-letter code) in the binding pocket of each of the NRPS modules as analyzed by NRPS-PKS software at http://www.nii.res.in/nrps-pks.html (Ansari et al., 2004).

[b] Amino acids (given in three-letter code) predicted to be activated by the respective NRPS module using the NRPSsp web-tool at http://www.nrpssp.com (Prieto et al., 2012).

variants, which are all in the fatty acid tail with the exception of glidobactin G, where a hydroxymethyl group replaces the methyl group attached to the ring structure (Oka et al., 1988a, 1988c; Terui et al., 1990). Because *glbH* was not needed for glidobactin A biosynthesis in *P. putida* carrying *plu1881-1877*, *glbH* would seem to play a role in a modification leading to a minor variant not present in *Pll*, and not in the hydroxylation of the lysine residue. Of course, it cannot be completely excluded that a gene in the *P. putida* genome substitutes for a *glbH* homologue, although BLAST analyses of the completely sequenced *P. putida* (and *Pll*) strains revealed no close homologue.

The *plu1881-1877* gene cluster also lacks a *glbE* homolog, which encodes an MbtH-like peptide. However, it has been shown that MbtH proteins encoded by different NRPS clusters are able to substitute for each other (Felnagle et al., 2010; Zhang et al., 2010b), and BLAST analysis reveals that both *Pll* and *P. putida* contain at least one MbtH-encoding gene in other NRPS gene clusters which may provide the function if needed.

5. The Endophytic Strain *Rhizobium sp.* AP16 Produces the Proteasome Inhibitor Syringolin A

Alexey Dudnik[1], Laurent Bigler[2], and Robert Dudler[1,3]

Submitted

[1] Institute of Plant Biology, University of Zurich, Zollikerstrasse 107, CH-8008 Zurich, Switzerland

[2] Institute of Organic Chemistry, University of Zurich, Winterthurerstrasse 190, CH-8057 Zurich, Switzerland

[3] Corresponding author

5.1. Abstract

Syringolin A, the product of a mixed non-ribosomal peptide synthetase/polyketide synthase encoded by the *syl* gene cluster, is a virulence factor secreted by certain *Pseudomonas syringae* strains. Together with the glidobactins produced by a number of β- and γ-proteobacterial human and animal pathogens, it belongs to the syrbactins, a structurally novel class of proteasome inhibitors. In plants, proteasome inhibition by syringolin A-producing *P. syringae* strains leads to the suppression of host defense pathways requiring proteasome activity, such as the ones mediated by salicylic acid and jasmonic acid. Here we report the discovery of a *syl*-like gene cluster with some unusual features in the α-proteobacterial endophyte *Rhizobium sp.* AP16 that encodes a putative syringolin A-like synthetase whose components share 55-65% sequence identity (72-79% similarity) at the amino acid level. As revealed by ANI (average nucleotide identity) calculation, this strain likely belongs to the same species as the biocontrol strain *R. rhizogenes* K84 (formely known as *Agrobacterium radiobacter* K84), which, however, carries a non-functional deletion remnant of the *syl*-like gene cluster. Here we present a functional analysis of the *syl*-like gene cluster of *Rhizobium sp.* AP16 and demonstrate that this endophyte synthesizes syringolin A and some related minor variants, suggesting that proteasome inhibition by syrbactin production not only can be important for pathogens but also for endophytic bacteria in the interaction with their hosts.

5.2. Introduction

Syringolin A was originally isolated from culture supernatants of the phytopathogenic γ-proteobacterium *Pseudomonas syringae* pv. *syringae* (*Psy*) B301D-R based on its ability to elicit defense responses and pathogen resistance in rice plants (Waspi, 2001; Wäspi et al., 1998b). It is a tripeptide derivative consisting an N-terminal valine and the two non-proteinogenic amino acids 3,4-dehydrolysine and 5-methyl-4-amino-2-hexenoic acid, whereby the first is N-acylated with an unusual ureido-valine moiety and the latter two form a 12-membered macrolactam ring (Figure 1A). Syringolin A is the major variant of a family of related compounds in which one or both valine residues can be replaced by isoleucine and/or 3,4-dehydrolysine can be substituted by lysine (Wäspi et al., 1999). Syringolin A was shown to be a virulence factor responsible for approximately 70% of the disease symptoms in the interaction of *Psy* strain B728a with its host plant *Phaseolus vulgaris* (bean) (Groll et al., 2008). The elucidation of the mode of action of syringolin A revealed that it

irreversibly inhibited all three proteolytic activities (i.e. the caspase-, trypsin-, and chymotrypsin-like activity) of the eukaryotic proteasome by covalent ether bond formation with the active-site N-terminal threonine residues of the catalytic β1, β2, and β5 subunits of the 20S core proteasome (Groll et al., 2008). Proteasome inhibition suppresses the action of many plant hormones, including defense reactions mediated by the important defense hormones jasmonic acid (JA) and salicylic acid (SA) (Dudler, 2013; Misas-Villamil et al., 2013; Schellenberg et al., 2010).

Syringolin A is structurally similar to glidobactin A and related variants (syn. cepafungins) that were isolated more than 20 years ago from the β-proteobacterial strain K481-B101 (ATCC 53080; DSM 7029; formerly misidentified as *Polyangium brachysporum*) and an undefined species related to *Burholderia cepacia* because of their antifungal and antitumor activities (Archer et al., 2010; Oka et al., 1988a, 1988b, 1988c; Terui et al., 1990). Similar to syringolin A, glidobactin A consist of a 12-membered ring structure and inhibits the eukaryotic proteasome by the same mechanism as syringolin A (Groll et al., 2008). In glidobactins, the ureido-valyl moiety is replaced by a fatty acid tail. Together, syringolin A and glidobactin A are the founding members of a novel class of proteasome inhibitors named syrbactins (Schellenberg et al., 2007).

Syringolin A and its variants are synthesized by a mixed non-ribosomal peptide synthetase/polyketide synthetase (NRPS/PKS) encoded by a gene cluster comprising the five open reading frames *sylA*-*sylE* (Figure 1B) (Amrein et al., 2004).Whereas *sylA* encodes a LuxR-type transcriptional activator of the *sylB* gene and the *sylCDE* operon (Ramel et al., 2012), *sylE* encodes a putative export facilitator involved in syringolin secretion. The *sylC* and *sylD* genes encode the NRPS/PKS responsible for syringolin biosynthesis, whereas *sylB* encodes a desaturase thought to mediate the conversion of lysine to 3,4-dehydrolysine in the ring structure. Based on the sequence and architecture of the *syl* gene cluster, an experimentally supported biosynthesis model of syringolin A was proposed which explains all structural features of the molecule (Amrein et al., 2004; Dudler, 2014; Imker et al., 2009; Ramel et al., 2009; Wuest et al., 2011). The syringolin variants are the result of incomplete lysine desaturation by SylB and a relaxed specificity of the SylC NRPS module, which, in addition to valine, activates also isoleucine, although with reduced efficiency (Imker et al., 2009; Wäspi et al., 1999).

Cloning of the glidobactin A synthetase revealed a gene cluster (*glb* genes) with high similarity in sequence and architecture to the *syl* gene cluster, allowing the postulation of an biosynthesis model analogous to the one for the syringolin variants (Schellenberg et al., 2007).

A search in genome sequence databases revealed intact *syl* gene clusters in the majority of sequenced strains belonging to phylogenetic group II of the *P. syringae* species complex (Baltrus et al., 2011; Hwang et al., 2005; Sarkar and Guttman, 2004), whereas, in addition to strain K481-B101, intact *glb*-like gene clusters were identified in the human pathogens *Burkholderia pseudomallei*, *B.oklahomensis* (β-proteobacteria), and *Photorhabdus asymbiotica* (γ-proteobacteria), as well as in the insect pathogen and nematode symbiont *Ph. luminescens* (Dudler, 2014; Schellenberg et al., 2007). Thus, these strains seem capable of synthesizing syrbactin-class proteasome inhibitors, and this was indeed shown for the genus *Photorhabdus*, which, in addition to glidobactin A (Dudnik et al., 2013; Theodore et al., 2012), also synthesizes the minor variant cepafungin I, the strongest proteasome inhibitor hitherto known (Stein et al., 2012).

Among recently deposited genome sequences we have discovered a *syl*-like gene cluster with unusual features in the genome of the α-proteobacterial strain *Rhizobium sp.* AP16, which was isolated from the root endosphere of poplar trees (*Populus deltoids,* eastern cottonwood) (Brown et al., 2012). Here we present a functional analysis of this gene cluster and its unique features and show that this endophyte synthesizes the proteasome inhibitor syringolin A.

5.3. Materials and methods

Bacterial strains and growth conditions. Bacterial strains and plasmids used in this study are listed in Table S1. *Escherichia coli* and *Pseudomonas syringae* were routinely grown in Luria-Bertani (LB) medium (tryptone 10 g/L, yeast extract 5 g/L, and NaCl 5 g/L). *Rhizozbium sp.* AP16 was cultivated in rhizobium medium (mannitol 10 g/L, K_2HPO_4 0.5 g/L, $MgSO_4$ 0.2 g/L, NaCL 0.1 g/L, yeast extract 1 g/L) or in SRM_{AF} medium (Mo and Gross, 1991) supplemented with 0.5 g/L yeast extract (SRM_{AFY}) unless stated otherwise. For solid media, 2% of agar was added. Antibiotics were used in the following concentrations: chloramphenicol: 20 μg/mL, gentamicin: 50 μg/mL, rifampicin: 50 μg/mL, and tetracycline: 15 μg/mL.

RNA extraction and RT-PCR analysis. RNA was extracted from bacteria grown as a lawn on SRM_{AFY} plates after 3 days in constant darkness using the RNeasy Protect Bacteria Mini Kit (QIAGEN, Hilden, Germany) with on-column DNAse treatment as described in the manufacturer's manual. Semi-quantitative reverse-transcriptase PCR (RT-PCR) was performed using QIAGEN OneStep RT-PCR Kit (QIAGEN, Hilden, Germany) as described in the manufacturer's manual.

Generation of markerless in-frame gene deletion mutants. 700 bp-long fragments of upstream and downstream sequences surrounding the target gene were amplified using the respective primer pairs (Supplementary table S2; primers P1 and P2 to amplify the upstream sequence, primers P3 and P4 amplify the downstream sequence). The primers were designed such that eighteen to thirty bases of the original gene sequence would be retained, including the start and the stop codons, in order to generate an in-frame deletion. The two DNA fragment were then joined by overlap extension PCR using primers P1 and P4. PCR product obtained was digested with the respective restriction enzymes, cloned into the suicide vectors pJQ200KSΔP*lac*, and transformed into *E. coli* XL-1 Blue. After verification by sequencing, the resulting construct was transformed into *E. coli* S17-1, and further mobilized into *Rhizobium sp.* AP16 by bi-parental mating. Single recombinants were selected on plates containing gentamicin, and, after verification by PCR, grown on plates with rhizobium medium supplemented with 10% sucrose in order to select double recombinants. Deletion of the target sequence was verified by PCR using the respective oriA and oriB primers, and sequence determination.

Isolation of syringolin from bacterial cultures and HPLC analysis. For *P. syringae,* liquid cultures (5 mL each) were grown overnight in LB medium supplemented with tetracycline (10 µg/mL) at 28°C on a shaker (220 rpm/min). The density was adjusted to an OD_{600} of 0.2 and 4-mL aliquots were used to inoculate 400-mL SRM_{AF} cultures in 1000-mL Erlenmeyer flasks. Flasks were further incubated at 20°C for 10 days under constant shaking (220 rpm). Cells were harvested by centrifugation and resuspended in 40 ml distilled water per flask. To isolate syringolin from *Rhizobium sp.* AP14 bacteria of 1.8 ml over-night cultures adjusted to an OD_{600} of 0.3 were collected, re-suspended in 200 µL distilled water (dH_2O), and plated out on SRM_{AFY} plates that were incubated for five days at 18^oC in constant darkness unless stated otherwise. In order to extract syringolin, cells were scraped off the plates, re-suspended in 40 mL distilled water per plate.

Resuspended cells were incubated on a shaker at 200 rpm for four hours. Cells were then harvested by centrifugation (3200x g, 10 min) and supernatants were transferred into fresh tubes supplemented with 2.5 g of Amberlite XAD-16 resin prepared according to the manufacturer (Sigma-Aldrich Chemie GmbH, Buchs, Switzerland). Tubes were incubated on a rotator at 30 rpm for two hours. Resin beads were then collected by centrifugation and washed three times in 40 mL dH_2O for 20 min on a rotator at 30 rpm. Beads were collected by sedimentation and washed twice with 20 mL 20% methanol for 20 minutes. After sedimentation of the beads, the supernatant was discarded and

syringolins were eluted with 900 µL of 70% isopropanol overnight at +4°C. Isopropanol was then removed and the beads were washed with another 900 µL of 70% isopropanol for 20 minutes at room temperature with shaking (200 rpm). The two eluates were combined, filtered through a Millex GP filter unit (0.22 µm pore size; Millipore, Molsheim, France), and 1 mL of it was evaporated to dryness using a Savant SpeedVac (Fisher Scientific Switzerland, Reinach, Switzerland). Dry material was re-suspended in 300 µL dH_2O, filtered again through a Millex GP filter unit, and stored at +4°C until further use.

HPLC analysis was performed as follows: Twenty µL of filtrate was injected into a 250 mm x 4.6 mm Nucleosil 100 C18, 5 µm column (Dr. Maisch GmbH, Ammerbuch, Germany) which was connected to UltiMate 3000 HPLC system (Thermo Fisher, Olten, Switzerland). Isocratic separation of components was performed with a 20% acetonitrile and 0.06% TFA in water at a flow rate of 1 mL/min.

Mass spectrometry and NMR. Low-resolution ESI mass spectra were measured with a quadrupole ion trap instrument (EsquireLC, Bruker Daltonik GmbH, Bremen, Germany). The solutions (about 0.1-1 µmol/ml in MeCN) were continuously introduced through the electrospray interface at a flow rate of 5 µl min-1. MS acquisitions were performed at normal resolution (0.6 u at half peak height), under ion charge control (ICC) conditions (10'000) in the mass range from *m/z* 100 to 2000. To get representative mass spectra, 8 scans were averaged.

High resolution (6’000) electrospray ionization mass spectrometry (HR-ESI-MS) was performed on a Finnigan MAT 900 (Thermo Finnigan, San Jose, CA, USA) double-focusing magnetic sector mass spectrometer. The samples were dissolved in CH_3OH (50 µM). PPG425 and PPG1000 (Aldrich, Steinheim, Germany) served for calibration.

All ^{1}H and ^{13}C NMR spectra were recorded at room temperature using a Bruker Avance 600 MHz instrument at 600 MHz (^{1}H) and 151 MHz (^{13}C). Samples (ca. 2 mg) were dissolved in d_6-DMSO. Spectra were calibrated with regard to proton chemical shift (δ = 0 ppm) and carbon chemical shift (δ = 0 ppm) of tetramethylsilane.

Phylogenetic tree generation. Nucleotide sequences of the four following housekeeping genes were extracted from NCBI Genbank (http://www.ncbi.nlm.nih.gov/genbank): DNA gyrase subunit B (*gyrB-1*), glyceraldehyde triphosphate dehydrogenase (*gap*), citrate synthase (*gltA*/*cisY*), and RNA polymerase sigma factor-70 (*rpoD*/*sigA*). Accession numbers of the included genomes are listed in

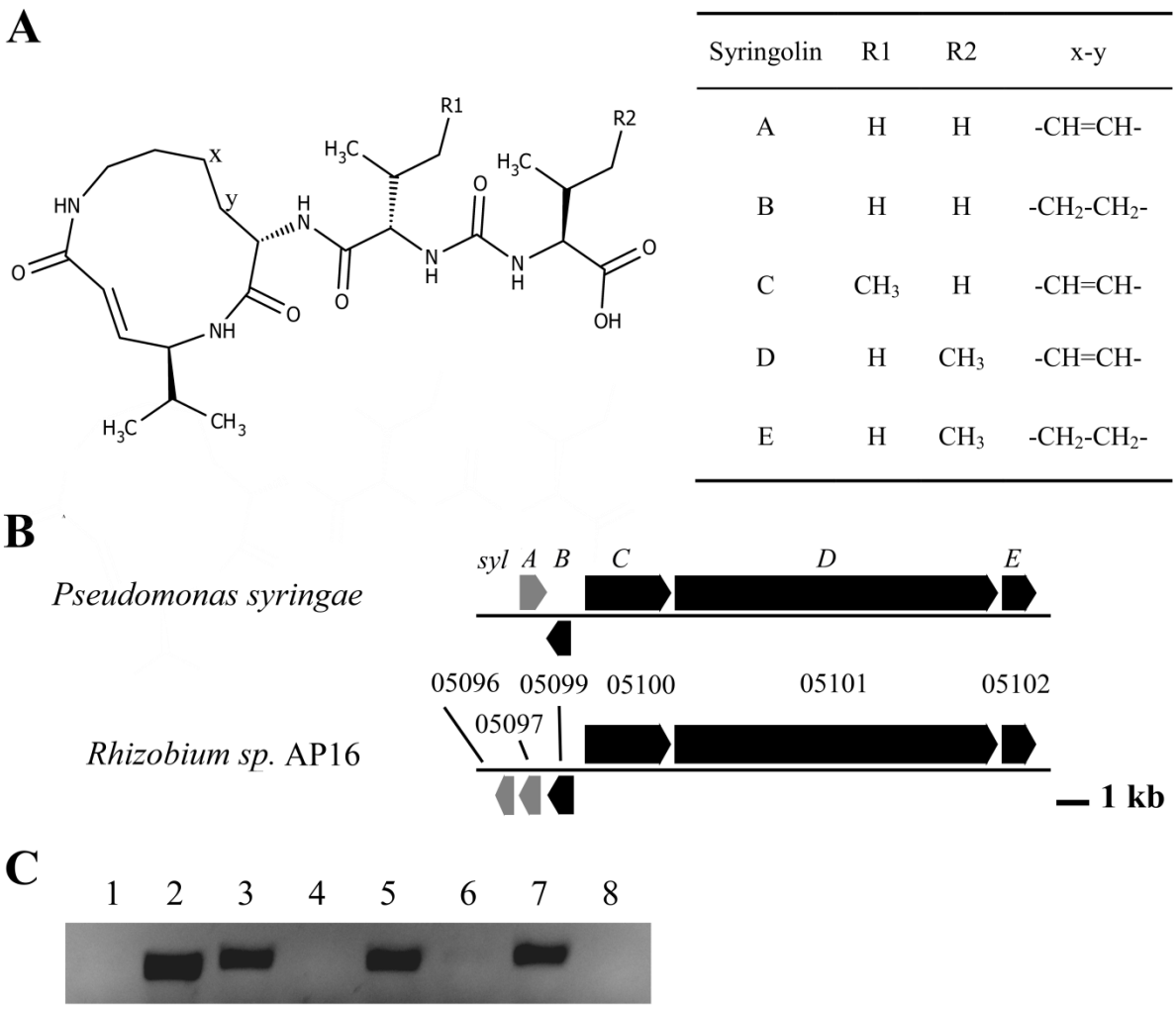
Supplementary table S3. Sequence alignment and phylogenetic tree construction was done using the MEGA 5.2.2 software (Tamura et al., 2011) using 1000 bootstrap cycles.

5.4. Results

5.4.1. The syl-like gene cluster of Rhizobium sp. AP16

BLAST searches in microbial genome databases at NCBI revealed the draft genome of *Rhizobium sp.* AP16 (AP16) to contain a region (NCBI accession NZ_AJVM01000074) encompassing a gene cluster with high similarity in sequence and architecture to the syringolin A synthetase-encoding genes of *Psy* B301D-R and a number of other *Psy* strains (Figure 1B) (Amrein et al., 2004). The proteins encoded by PMI03_05099-PMI03_5102 exhibit 55-65% sequence identity (72-79% similarity) to the products of the *sylB-sylE* genes that are responsible for biosynthesis and export of syringolin A in *Psy* strains. The AP16 gene cluster lacks a homolog of the *sylA* gene encoding a transcriptional activator of the *sylB* gene and the *sylCDE* operon in *Psy* B301D-R, suggesting that transcriptional regulation is not conserved. In addition to these genes, the AP16 gene cluster contains three open reading frames (ORFs) (PMI03_05098- PMI03_05096) on the bottom strand downstream of the *sylB* homolog PMI03_05099 which are not present in *Psy syl* gene clusters (Figure 1B). The relatively close spacing of these ORFs (78 bp between PMI03_05099 and PMI03_05097; 69 bp between PMI03_05097 and PMI03_05096) may suggest that they are transcribed as an operon together with PMI03_05099, which would point to a function of these additional genes in the biosynthesis of the putative product controlled by the gene cluster. The annotation of PMI03_05098 appears doubtful as it is predicted to encode a short peptide of 35 amino acids with no similarity to known proteins and overlaps 25 nucleotides with PMI03_05097, which encodes a type II (stand-alone) thioesterase. Type II thioesterases are known to release stalled NRPS modules that have been loaded with a non-extendable moiety (Schwarzer et al., 2002; Yeh et al., 2004). The third ORF, PMI03_05096, encodes a hypothetical protein of unknown function belonging to 5-carboxymethyl-2-hydroxymuconate isomerase (CHMI) superfamily of proteins (cl01268). Certain members of this superfamily catalyze the isomerization of unsaturated ketones (Poelarends et al., 2008).

Because of the highly conserved domain structure and amino acid specificity as revealed by the ’10-amino-acid codes’ (Stachelhaus et al., 1999) of the NRPS modules in the AP16 SylC (PMI03_05100) and SylD (PMI03_05101) homologs, the AP16 gene cluster is predicted to control



Syringolin	R1	R2	x-y
A	H	H	-CH=CH-
B	H	H	$-CH_2-CH_2-$
C	CH_3	H	-CH=CH-
D	H	CH_3	-CH=CH-
E	H	CH_3	$-CH_2-CH_2-$

Figure 1. (A) Chemical structure of syringolin variants. The core structure is given on the left and the modifications present in all accounted variants are listed on the right (B) Structures of syringolin biosynthesis gene clusters from *P. syringae* pv. *syringae* B301D-R and *Rhizobium sp.* AP16. Homologous ORFs are shown in black while ORFs that are unique to one of the gene clusters are depicted in grey. Gene names are positioned atop of the corresponding ORFs. (C) Expression of selected genes from the PMI03_05096-05102 gene cluster on SRM_{AFY}-agar at 18°C as detected by semi-quantitative reverse transcriptase (RT)-PCR. Label "inter-genic" indicates when primers originate in two adjacent genes, thus covering the corresponding inter-genic region. The following genes were assayed: (1) PMI03_05096, (2) PMI03_05097, (3) PMI03_05099, (4) PMI03_05097-05099 inter-genic, (5) PMI03_05100, (6) PMI03_05100-5101 inter-genic, (7) PMI03_05102, (8) no-RT control.

the biosynthesis of a syringolin A-like proteasome inhibitor. The closely spaced additional genes downstream of the *sylB* homolog PMI03_05099 and the unusual tandem thiolation domains in the SylC homolog PMI03_05100 might be involved in decoration and/or biosynthesis efficiency of the putative product.

5.4.2. Rhizobium sp. AP16 produces syringolin A and related variants

We sought to detect transcripts corresponding to the *syl*-like gene cluster of the AP16 strain under various growth conditions. The strain grew reasonably well on plates with rhizobium medium both at 28°C and 18°C, but hardly on LB plates. Best growth was obtained on SRM_{AF} medium supplemented with 0.5 g/L yeast extract (SRM_{AFY}). RNA was extracted from bacteria collected from rhizobium medium or SRM_{AFY} plates after incubation at 18°C or 28°C for three days in the dark and used as a template for RT-PCR using primers specific for the individual ORFs (PMI03_05096 through PMI03_05102, except PMI03_05098, which was not considered further). None of these genes was detectably expressed in bacteria grown on plates with either medium at 28°C. In contrast, growth on SRM_{AFY} plates at 18°C revealed transcripts corresponding to PMI03_05097 to PMI03_05102, whereas PMI03_05096 transcripts were not detected (Figure 1C). On rhizobium medium plates, transcripts were hardly detectable. Therefore, SRM_{AFY} plates were used for all further experiments. Amplification by RT-PCR of the intergenic region between PMI03_05100 and PMI03_05101 was not successful (Figure 1C), indicating that these genes, unlike their homologs in syringolin A-producing *Psy* strains, were not transcribed as an operon (Ramel et al., 2012). Similarly, the intergenic region between PMI03_05097 and PMI03_05099 was also not detected in the extracted RNA.

To detect the putative product of the AP16 *syl*-like gene cluster, bacteria were grown as a lawn on SRM_{AFY} plates at 18°C in the dark for five days, scraped off the plates and subjected to our standard syringolin extraction and detection procedure (see methods section). Samples were subjected to HPLC analysis. For comparison of HPLC profiles, a PMI03_05100 (*sylC* homolog) in-frame deletion mutant was constructed (AP16_ΔSylC) and analyzed in parallel. As shown in Figure 2, several peaks were detected in wild-type extracts that were absent in the AP16_Δ*sylC* mutant. Comparison with HPLC profile of extracts from *Psy* B301D-R suggested that the major peaks A and D in *Rhizobium sp.* AP16 extracts corresponded to syringolin A and syringolin D of B301D-R (Wäspi et al., 1999). Peak A from strain AP16 was collected and subjected to LC-HR-ESI-MS, which revealed a quasi-molecular ion with an *m/z* of 494.29747, corresponding to the formula $C_{24}H_{40}N_5O_6$ (hydrogen adduct, *m/z* 494.29731, SD 0.32 ppm), which is identical to syringolin A. The ^{1}H and ^{13}C NMR data of peak A were also essentially identical to the ones of syringolin A isolated from *Psy* B301D-R (Supplementary table S4). We conclude from these results that *Rhizobium sp.* AP16 synthesizes the proteasome inhibitor syringolin A under the culture conditions

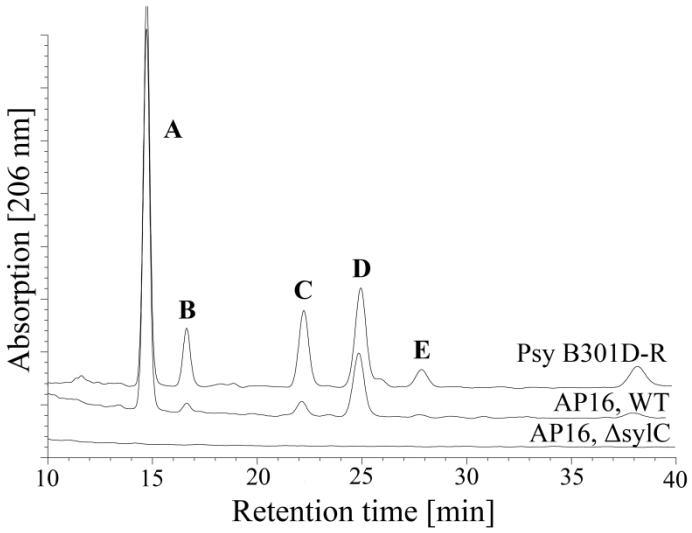


Figure 2. Aligned HPLC profiles of *P. syringae* pv. *syringae* B301D-R (pOEAC), *Rhizobium sp*. AP16 wild type (WT) and PMI03_05100 deletion mutant (Δ*sylC*).

used. Interestingly, when AP16 was grown on plates in the light (16h light/8h dark), syringolin A production was nearly abolished (Supplementary figure S1). Syringolin A was also not produced when AP16 was grown in liquid culture, irrespective of the light conditions.

ESI-MS of the second largest AP16 peak, peak D, revealed a quasi-molecular ion with an *m/z* of 508.2, which is identical to the hydrogen adduct of syringolin D, thus confirming the tentative identification given above. In syringolin D, ureido-valine is replaced by ureido-isoleucine (Wäspi et al., 1999). The comparison of the HPLC profiles showed that AP16 produced significantly less syringolin B, C, and E relative to syringolin A than observed in *Psy* B301D-R (Table 1). In syringolin B, the 3,4-dehydrolysine moiety in the ring structure is replaced by lysine, which, according to the biosynthesis model, results from incomplete lysine desaturation by the *sylB* gene product, most likely before incorporation (Amrein et al., 2004; Wuest et al., 2011). Thus, lysine desaturation seems considerably more efficient in AP16 as compared to *Psy* B301D-R. This explains also the reduced abundance of syringolin E in AP16, as this variant combines the saturated lysine with the ureido-isoleucine (Table 1). In addition, syringolin C, in which the valine attached to the ring structure is replaced by isoleucine, is less abundant in AP 16 as compared to *Psy* B301D-R (Figure 2 and Table 1).

In order to test whether PMI03_05097 and/or PMI03_05096, which are located downstream of the *sylB* homolog PMI03_05099 on the same DNA strand (Figure 1B), had a function in syringolin production, in particular in the apparently more efficient desaturation of lysine to 3,4-dehydrolysine

Table 1. Production of syringolin variants in relation to syringolin A

Strain	Syringolin[a]				
	A	B	C	D	E
B301D-R	100	17.7 ± 2.0	30.2 ± 2.0	37.9 ± 2.4	9.2 ± 1.3
AP16	100	2.6 ± 0.3	5.3 ± 0.4	27.6 ± 2.0	0.5 ± 0.1
Δ05097	100	3.0 ± 1.0	6.3 ± 2.2	24.8 ± 1.8	1.1 ± 0.9
Δ05096/97	100	2.7 ± 1.0	6.6 ± 1.3	27.4 ± 0.6	0.6 ± 0.3
ΔPCP2	100	1.5 ± 0.6	5.7 ± 0.5	25.6 ± 0.3	0.6 ± 0.1

[a] Average relative peak area ± standard deviation with respect to syringolin A. Data were derived from HPLC profiles (n=3).

in strain AP16 as compared to *Psy* B301D-R, in-frame mutants carrying in-frame deletions of PMI03_05097 (Δ05097) or both of PMI03_05096 and PMI03_05097 (Δ05096-05097) were constructed. Comparison of HPLC profiles of these two mutants and the wild type AP16 strain did not exhibit significant differences with regard to relative variant abundances (Figure 3A and Table 1), thus suggesting that PMI03_05096 and PMI03_05097 did not cause the different variant abundances observed in *Psy* B301D-R and AP16. However, both mutants reproducibly produced reduced amounts of syringolin A: In Δ05097 and Δ05096-05097, syringolin A levels reached 59 ± 0.4% (p=0.002, tTest, n=3) and 49 ± 11% (p=0.004, tTest, n=3) of the values observed in the wild type AP16 strain, respectively, suggesting that PMI03_05097 gene enhances biosynthesis efficiency of syringolin.

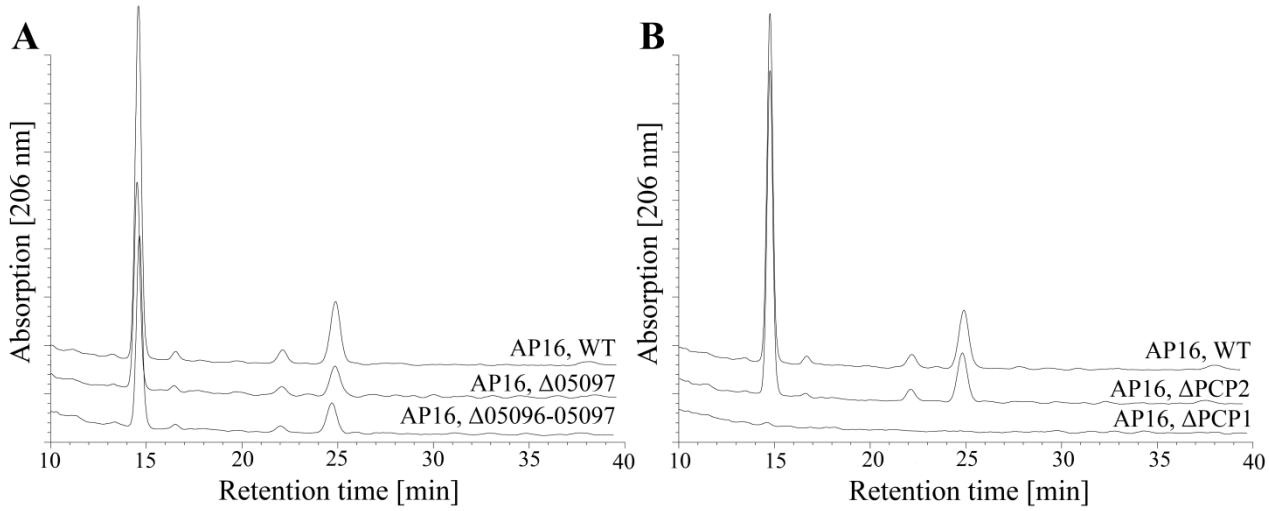


Figure 3. (A) HPLC profiles of *Rhizobium sp.* AP16 wild type (WT), deletion mutant of PMI03_05097 (Δ05097) and double deletion mutant of PMI03_05096 and PMI03_05097 (Δ05096-05097). (B) Alignment of the HPLC profiles of *Rhizobium sp*. AP16 wild type (WT) with deletion mutants of the thiolation domains 1 and 2 in PMI03_05100, (ΔPCP1) and (ΔPCP2) respectively. Note that in both cases the chromatograms were drawn to scale, i.e. peak sizes are directly comparable between the different graphs.

5.4.3. The AP16 SylC homolog PMI03_05100 exhibits tandem thiolation domains

Another unusual feature of the *syl*-like gene cluster of AP16 is that the NRPS module encoded by PMI03_05100, the AP16 *sylC* homolog, contains two tandemly arranged thiolation domains (also known as peptide carrier proteins or PCPs). Tandem thiolation domains have been described as acyl carrier proteins (ACPs) in a number of PKS modules and polyunsaturated fatty acid synthetases where they have been implicated in the enhancement of product formation (El-Sayed et al., 2003; Gu et al., 2011; Jiang et al., 2009).

To determine whether the tandem thiolation domains had an influence on syringolin production, in-frame deletion mutants of PMI03_05100 were generated that lacked either the first (ΔPCP1) or the second (ΔPCP2) thiolation domain. HPLC analyses of the mutants and the wild type revealed that ΔPCP2 produced 107 ± 12% of the wild type AP16 (p=0.43, tTest, n=3), which is not significantly different from the wild-type amount. Relative variant abundances in ΔPCP2 were also the same as in the wild type (Figure 3B and Table 1) In contrast, syringolin biosynthesis was completely abolished in ΔPCP1 (Figure 3B). Thus, a single PCP (PCP1) appears to be sufficient for full syringolin A biosynthesis under the culture conditions used. The syringolin-negative phenotype of ΔPCP1 could be due to improper folding of this internal deletion protein, or it may indicate that PCP2 is unable to substitute for PCP1.

5.4.4. Rhizobium sp. AP16 is closely related Rhizobium rhizogenes K84

Rhizobium rhizogenes K84 (formerly *Agrobacterium radiobacter* K84 (Velázquez et al., 2010)) has been commercially used as a biocontrol agent of crown gall disease for decades (Farrand, 1990). As visualized in the phylogenetic tree shown in Figure 4 *Rhizobium sp.* AP16 is closely related to *R. rhizogenes* K84. In fact, direct comparison of the complete draft genome of *Rhizobium sp.* AP16 (accession NZ_AJVM00000000; 96 contigs) with the complete genome sequence of *R. rhizogenes* K84 (accessions NC_011985 and NC_011983 for chromosomes 1 and 2) using the ANI (average nucleotide identity) calculator (http://enve-omics.ce.gatech.edu/ani/) yielded an identity value of 98.5%, strongly suggesting that the two strains belong to the same species (Goris et al., 2007). This conclusion was corroborated by a comparison using the Genome-to-Genome Distance Calculator GGDC 2.0 program GBDP2_BLASTPLUS (http://ggdc.dsmz.de/distcalc2.php; (Goris et al., 2007; Meier-Kolthoff et al., 2013)), which assigned the two genomes to the same species with a probability of 98.97%.

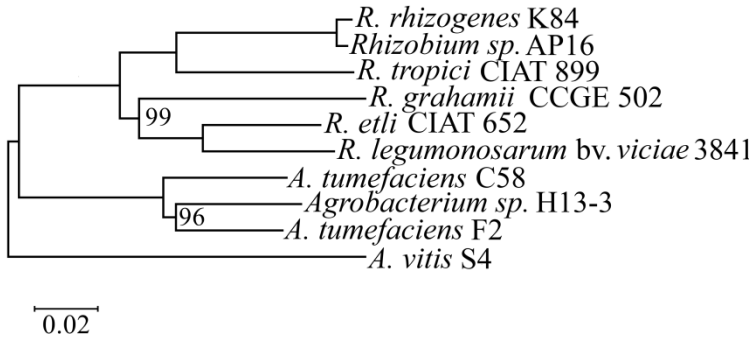


Figure 4. Neighbor-joining tree illustrating the phylogenetic position of *Rhizobium sp*. AP16. The tree was constructed based on nucleotide sequences of four housekeeping genes. The scale bar indicates the number of nucleotide changes per site. Bootstraping-based probabilities are displayed only at nodes where the values are different than 100.

Interestingly, *R. rhizogenes* K84 carries a deletion remnant of the *syl* gene cluster on chromosome 2 consisting of the sequence encoding the last 102 amino acids of the PMI03_05101 (*sylD*) homolog and the complete PMI03_05102 (*sylE*) homolog (annotated as YP_002541055 in *R. rhizogenes* K84) (Dudler, 2013). The deletion encompasses the rest of the *syl* gene cluster and the part of the PMI03_05095 homolog encoding the N-terminal 43 amino acids.

5.5. Discussion

We have demonstrated that the AP16 *syl*-like gene cluster indeed governs the biosynthesis of the proteasome inhibitor syringolin A and some related variants. In comparison to *Psy* B301D-R, AP16 produces approximately 10-fold less syringolin B and only vanishing amounts of syringolin E in relation to syringolin A. SylB, which is homologous to dihydrorhizobitoxin desaturase (Yasuta et al., 2001), is thought to be responsible for the desaturation of lysine to 3,4-dehydrolysine (Amrein et al., 2004). This reaction probably occurs before incorporation into the peptide structure by the first NRPS module of SylD, because its activation domain adenylates 3,4-dehydrolysine much faster and more efficiently than lysine (Wuest et al., 2011). The low abundance of syringolin B and E in AP16 shows that lysine is incorporated into the peptide at a much lower rate than in *Psy* B301D-R. This difference seems to be due to the intrinsic properties of the corresponding NRPS modules of the two species. The alternative hypothesis we initially entertained, namely that the stand-alone thioesterase encoded by PMI03_05097 might lower the incorporation rate of the non-preferred lysine residue, had to be discarded because the deletion of PMI03_05097 had no influence on variety ratios (Table 1).

Syringolin C and D are isomers differing in the substitution of the ring-proximal or the distal valine residue by isoleucine (Figure 1A) (Wäspi et al., 1999). Whereas the abundance ratio of these two variants in *Psy* B301D-R is 0.8, it is 0.2 in AP16 (Table 1), i.e. syringolin C has a four-fold lower abundance in the latter strain. *Psy* SylC was experimentally shown as the NRPS starter module to activate valine, and to a lesser extent, isoleucine and to form an ureido-dipeptide, whereby the ureido carbonyl group is derived from hydrogen carbonate/carbon dioxide (Imker et al., 2009; Ramel et al., 2009). In contrast to *Psy*, in AP 16, valine is much less likely to be substituted by isoleucine at the ring-proximal position than at the distal position. This appears to be an intrinsic property of the SylC homolog encoded by PMI03_05100 and is not due to the unusual tandem PCP domains, because the ΔPCP2 mutant and the wild type have very similar absolute variant abundances (Figure 3B and Table 1). Thus, the second PCP in PMI03_05100 seems to be redundant for any aspect of syringolin biosynthesis under the experimental conditions used. Apart from the protein encoded by PMI03_05100, only one other NRPS module exhibiting tandem PCP domains was identified by a search through the NCBI RefSeq protein sequence database (accession WP_007460164 from *Streptomyces sp*. W007), which, however, remains uncharacterized. As mentioned above, tandem thiolation domains (ACPs) have been described in some type I PKS modules as well as in polyunsaturated fatty acid (PUFA) synthetases. In some cases, they have been demonstrated to improve product yield by working either serially or in parallel (El-Sayed et al., 2003; Gu et al., 2011; Jiang et al., 2009).

AP16 has been isolated from the endosphere of poplar roots. Many bacteria have been found to entertain endophytic relationships with their host plants (van der Lelie et al., 2009; Reinhold-Hurek and Hurek, 2011; Rosenblueth and Martínez-Romero, 2006). Plants have been shown to respond with local defense reactions to colonization by endophytes (Bordiec et al., 2011; Conn et al., 2008; Miché et al., 2006). It also results in elevated levels of SA. Compatible rhizobia produce specific nodulation (Nod) factors, which in turn suppress SA production, thus down-regulating defense responses (Mabood and Smith, 2007). Based on our analysis, the genome of AP16 lacks any known *nod* genes. In addition, despite having a seemingly complete type III secretion system (contig AJVM01000009), we did not detect any known rhizobial or *Pseudomonas* type III effector proteins, which are normally involved in suppression of host immune responses (Jones and Dangl, 2006). Thus, it is tempting to speculate that in AP16, syringolin A can substitute for type III effectors or Nod factors with regard to host defense suppression during the colonization of poplar roots. Further experiments are needed to clarify this point.

5.6. Supplementary figures and tables

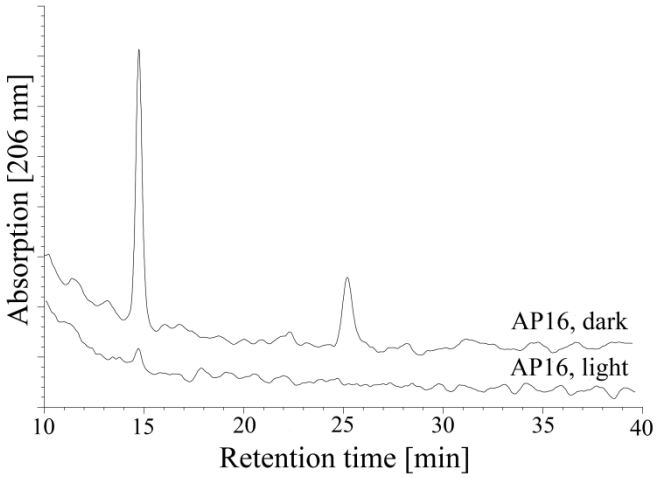


Figure S1. HPLC profiles of *Rhizobium sp*. AP16 wild type grown under constant darkness (dark) or long-day conditions (16h light/8h darkness). Note that in both cases the chromatograms were drawn to scale, i.e. peak sizes are directly comparable between the different graphs.

Table S1. Bacterial strain and plasmids used in the study

Strain or Plasmid	Genotype, relevant characteristics	Source or Reference
Rhizobium sp. AP16	Wild type strain, Chl^R	(Brown et al., 2012)
AP16, Δ*sylC*	Deletion mutant of PMI03_05100 (*sylC*)	This study
AP16, Δ05097	Deletion mutant of PMI03_05097	This study
AP16, Δ05096-05097	Double deletion mutant of PMI03_05096 and PMI03_05097	This study
AP16, ΔPCP-1	Deletion mutant of PCP-1 site of PMI03_05100 (*sylC*)	This study
AP16, ΔPCP-2	Deletion mutant of PCP-2 site of PMI03_05100 (*sylC*)	This study
P. syringae pv. *syringae* B301D-R (pOEAC)	Syringolin-producing strain; SylA over-expression/sylC promoter-reporter gene fusion	(Ramel et al., 2012)
Escherichia coli		
XL-1 Blue	*endA*1 *gyrA*96(nal^R) *thi*-1 *recA*1 *relA*1 *lac glnV*44 F'[::Tn10 $proAB^+$ $lacI^q$ Δ(*lacZ*)M15] *hsdR*17(r_K^- m_K^+)	(Bullock et al., 1987)
S17-1	*thi pro hsdR recA*; chromosomal RP4 [Tra+ Tc^S Km^S Ap^S]	(Simon et al., 1983)
Plasmids		
pJQ200KSΔP*lac*	Suicide vector; p15A ori; $Gent^R$, SacB, Δ*lac* promoter	This study

Table S2. List of primers used in the study

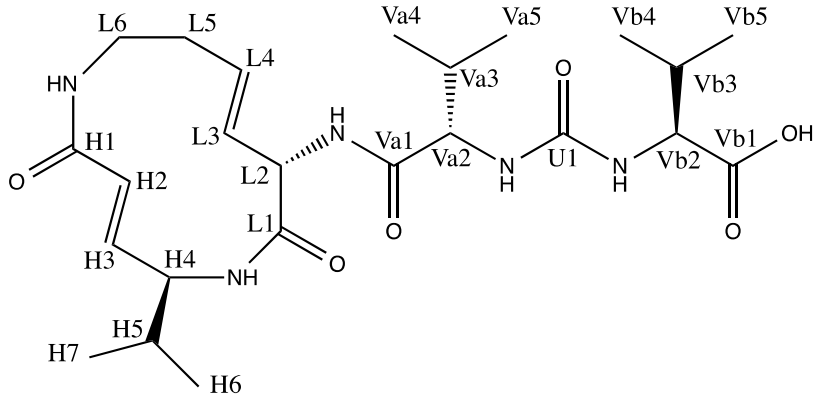
Primer	Sequence
PMI03-5097_rt_f	ATC GAA ATA TCG GGC ATA CG
PMI03-5097_rt_r	TTT CCA CCA TTT CGA AGT CC
PMI03-5100_rt_f	GTG TCA CTG GGC TCG ATT TC
PMI03-5100_rt_r	AAC CCG ACC TGG ATA TAG GG
PMI03-5102_rt_f	CAG TTT GTA CGC GCT GTT TC
PMI03-5102_rt_r	GTC AAG GCG TAG GAA CCA AG
AP16-5099_rt_f	CGG TGT TTC AGG ACT TCG AT
AP16-5099_rt_r	GAG TTC GTC GTC TTC CTT GC
AP16-5099-97_rt_f	CCC TAC TGG AAC CTG CAG AA
AP16-5099-97_rt_r	CGT ATG CCC GAT ATT TCG AT
AP16-5097_rt_f	TTT GCA TTC TTC GGT CAC AG
AP16-5097_rt_r	TAT GTG CGT GTC CGT TCT TC
AP16-5096_rt_f	CCG CTG TCA GAA TAG AAA CG
AP16-5096_rt_r	ATC CAT CGG TCT TGA AAT CG
AP16-5100-01_rt_f	ACG TCC TCA AGC TCG ACA AG
AP16-5100-01_rt_r	CGA ACG GAA CAT TGT ACG TG
5097_del_P1	ACT AGG TCT AGA AGG ACC GGT TTG AAT TCG TC
5097_del_P2	CGC TCA ATG CGC CGA ACT CAT GAT ATT TCC TTT C
5097_del_P3	ATC ATG AGT TCG GCG CAT TGA GCG CGA TTT G
5097 _del_P4	ATC AGT CGA CGC TTA TTG CCT ATG GGT TGC
5097_del_oriA	CGC TTC ATC GGC CAA TAC
5097_del_oriB	GGC TGA ATT CGG AGA ATA CC
5096-97_del_P1	ACT AGG TCT AGA TTC GTC TTC AAG CAC CTT CTC
5096-97_del_P2	CGA TCA TCC ATC GGA CGA ACT CAT GAT ATT TCC TTT C
5096-97_del_P3	ATG AGT TCG TCC GAT GGA TGA TCG CCG CAA AAC
5096-97_del_P4	ATC AGT CGA CTA CGG CTA TCC TGA CGC TTG
5096-97_del_oriA	ACA CGT CAA GAT GCA TCA CC
5096-97_del_oriB	GGG CGT GGA TAT CAG TCT TC
5100_del_P1	ACT AGG TCT AGA GAC CTT GAG CAG GAA TAC GG
5100_del_P2	GCT CAG AGA TGC GCG ACA AGA TAG CTC ATG AAG CAC TC
5100_del_P3	TGA GCT ATC TTG TCG CGC ATC TCT GAG CGC TGC
5100_del_P4	ATC AGT CGA CCC TAT CGC CAG ATC CTC ATC
5100_del_oriA	AAC GGA GAG CAG CTT GAG TG
5100_del_oriB	CCA ACC ATG TAC AGC GAC AC
PPI1_del_P1	TAT GGA TCC AGC AGG TAA GCA TCC TGT GG
PPI1_del_P2	AGG CAA CGA TTG GTA GAC GTT CAG GGG TAC CAT G
PPI1_del_P3	CTG AAC GTC TAC CAA TCG TTG CCT GAC CAT AC
PPI1_del_P4	ATC AGT CGA CTG CTG CAC GAC ATC AAG TG
PPI1_del_oriA	GTT CGA CGC ATC CAC ATT C

PPI1_del_oriB	GAC AGG TCC ACC TGG TTC AG
PPI2_del_P1	AAG GAA AAA AGC GGC CGC ACT TGA AAC GCC CTG AAC TG
PPI2_del_P2	ATC TGT GAT AGC CGA AGT ATG GTC AGG CAA CG
PPI2_del_P3	GAC CAT ACT TCG GCT ATC ACA GAT GGC CAC GAC G
PPI2_del_P4	TAT GGA TCC CCT ATC GCC AGA TCC TCA TC
PPI2_del_oriA	ATC GCG ATC TCT ACG TGG TC
PPI2_del_oriB	CCC TTG AAT TGT CGT GTG G

Table S3. Strains used from construction of the phylogenetic tree and the corresponding accession numbers

Strain	**Accession Number**
Rhizobium rhizogenes K84	CP000628 - CP000632
Agrobacterium sp. H13-3	CP002248 - CP002250
Agrobacterium tumefaciens F2	AFSD00000000
Agrobacterium vitis S4	CP000633 - CP000639
Agrobacterium fabrum C58	AE007869 - AE007872
Rhizobium sp. AP16	AJVM00000000
Rhizobium leguminosarum bv. *viciae* 3841	AM236080 - AM236086
Rhizobium etli CIAT 652	CP001074 - CP001077
Rhizobium tropici CIAT 899	CP004015 - CP004018
Rhizobium grahamii CCGE 502	AEYE00000000

Table S4. ^{1}H- and ^{13}C-NMR (DMSO) of Syringolin A, incl. chemical shift [ppm] and interpretation



	^{1}H (Syringolin A)		**^{13}C (Syringolin A)**	
	AP16	**B301D-R$_{ref}$**	**AP16**	**B301D-R$_{ref}$**
H1	–	–	166.2	166.2
H2	6.10	6.10	121.4	121.4
H3	6.68	6.68	143.2	143.2
H4	4.07	4.07	55.4	55.4
H5	1.73	1.73	31.4	31.3
H6	0.94	0.94	19.6	19.6
H7	0.91	0.90	19.1	19.1
H4NH	8.09	8.08	–	–
L1	–	–	168.7	168.7
L2	4.87	4.86	53.4	53.4
L3	5.39	5.40	126.3	126.3
L4	5.59	5.59	132.7	132.7
L5	2.27&1.96	2.27&1.95	34.9	34.9
L6	3.17	3.17	42.4	42.4
L6NH	7.45	7.45	–	–
L2NH	7.78	7.78	–	–
Va1	–	–	171.8	171.8
Va2	3.95	3.95	57.9	57.9
Va3	1.95	1.95	30.3	30.3
Va4&Va5	0.77	0.77	17.6	17.6
	0.80	0.80	19.4	19.3
Va2NH	6.41	6.40	–	–
Vb1	–	–	174.6	174.6
Vb2	3.62	3.64	59.7	59.6
Vb3	1.95	1.95	31.4	31.3
Vb4&Vb5	0.75	0.75	18.2	18.2
	0.76	0.76	19.7	19.7
Vb2NH	6.06	6.07	–	–
U1	–	–	157.7	157.7

All ^{1}H and ^{13}C NMR spectra were recorded using a Bruker Avance 600 MHz (^{1}H) and 151 MHz (^{13}C) at room temperature. Samples (ca. 2 mg) were dissolved in d_6-DMSO. Chemical shifts (δ-values) are reported in ppm, spectra were calibrated with regards to proton chemical shift (δ = 0 ppm) and carbon chemical shift (0 ppm) of tetramethylsilane.

6. Virulence Determinants of *Pseudomonas syringae* Strains in the Context of a Small Type III Effector Repertoire

Alexey Dudnik[1] and Robert Dudler[1,2]

Submitted

[1] Institute of Plant Biology, University of Zurich, Zollikerstrasse 107, CH-8008 Zurich, Switzerland

[2] Corresponding author

6.1. Abstract

Pseudomonas syringae is pathogenic to a large number of plant species. For host colonization and disease progression, the strains of this bacterium utilize an array of type III-secreted effector proteins and other virulence factors, including small secreted molecules such as the phytotoxin syringolin A. Syringolin A is a non-ribosomally synthesized tetrapeptide derivative that inhibits the eukaryotic proteasome. In strains colonizing dicotyledonous plants, the compound was demonstrated to suppress the salicylic-acid-dependent defense pathway and stomatal immunity, thereby enhancing virulence. Here, we analyze virulence factors of three strains colonizing wheat (*Triticum aestivum*): *P. syringae* pv. s*yringae* strains B64 and SM, as well as *P. syringae* BRIP34876. These strains have a relatively small repertoire of only seven to eleven type III secreted effectors (T3Es) and differ in their capacity to produce syrinolin A. We analyze the contribution of various virulence factors in the context of a small T3E repertoire. We demonstrate that syringolin A influences disease symptom development in a dosage-dependent manner. Despite the small set of T3E-encoding genes, the type III secretion system remains the key pathogenicity determinant, and phenotypic effects of deleting effector genes become apparent only when more than one is removed. In addition, we demonstrate that the type VI secretion systems in these strains are involved in competition against other bacteria, but do not play a role in virulence.

6.2. Introduction

The genus *Pseudomonas* of the Gram-negative γ-proteobacteria includes a number of species which directly or indirectly influence our everyday life. Members of the genus are metabolically versatile and are associated with various ecological niches and life styles (Palleroni, 2005b). It includes soil bacteria with biocontrol properties (*P. fluorescens*, *P. putida*, and *P. chlororaphis*) (Chin-A-Woeng et al., 2000; Palleroni, 2005b; Saxena et al., 2000), opportunistic human pathogens such as *P. aeruginosa* and *P. stutzeri* (Ceri et al., 2010; Gellatly and Hancock, 2013), as well as the plant-pathogenic species *P. fuscovaginae*, *P. marginalis*, and *P. syringae* (Höfte and De Vos, 2006; O'Brien et al., 2011; Palleroni, 2005b). Besides of *P. aeruginosa*, *P. syringae* is the second best-studied member of this genus, mainly due to its economic impact, but also because of its high genetic and metabolic flexibility that results in a variety of induced disease types and colonized hosts species (Mansfield et al., 2012). The species also serves as a model for plant-pathogen

interaction research, mainly with regard to the type III secretion system (T3SS) and effector function (Lindeberg et al., 2012), plant defense signaling, and gene regulation (Gimenez-Ibanez and Rathjen, 2010; Qi and Innes, 2013).

The T3SS is a complex structure which is used by a number of animal and plant pathogens to deliver so-called effector proteins into cells of their host (Büttner, 2012; Galán and Wolf-Watz, 2006). The type III-translocated effectors (T3Es) are in turn modulating the target cells in a variety of ways. For example, they can suppress defense and other signaling cascades, modify of cytoskeleton structure and gene transcription, or interfere with intracellular trafficking (Cunnac et al., 2009; Lee et al., 2013). The majority of characterized T3Es from *P. syringae* were demonstrated to be involved in the suppression of plant immune responses (Deslandes and Rivas, 2012; Jones and Dangl, 2006; Zhou and Chai, 2008). This is achieved in a variety of ways, and several effectors were demonstrated to degrade components of a defense signaling pathway either directly (Mackey et al., 2003; Zhang et al., 2010a) or by ubiquitylation (Göhre et al., 2008). Other T3Es are known to interfere with signaling either by covalently modifying one of the mitogen-activated protein (MAP) kinases (Wang et al., 2010), or by inhibiting its kinase activity (Xiang et al., 2008). Effectors were also shown to inhibit callose deposition (DebRoy et al., 2004) and production of reactive oxygen species (Rodríguez-Herva et al., 2012), both of which are activated upon pathogen perception. It should be noted that it is not uncommon for different effectors to target the same defense pathway, or even the same component within the pathway, thus having a redundant function (Cunnac et al., 2011; Deslandes and Rivas, 2012; Lee et al., 2013). The T3E composition is the key element determining host specificity (Lindeberg et al., 2009). Nonetheless, due to functional redundancy, strains isolated from the same host often show variability in their effector sets (Almeida et al., 2009; O'Brien et al., 2012; Qi et al., 2011).

Aside of T3Es, *P. syringae* strains often produce other substances which promote virulence. One such group of compounds comprises the phytotoxins (Bender et al., 1999). Based on their targets and mode of action the phytotoxins can be subdivided into several classes. One group consists of the chlorosis-inducing antimetabolite toxins which inhibit various nitrogen metabolism enzymes. It includes phaseolotoxin, tabtoxin, mangotoxin, and several other uncharacterized toxins (Arrebola et al., 2011; Bender et al., 1999). Another group encompasses cyclized lipopeptides syringomycin and syringopeptin, which are thought to induce ion leakage through insertion into the host cell membrane (Bender et al., 1999). A third type of phytotoxin is tagetitoxin, which inhibits bacterial,

chloroplast, and some eukaryotic RNA polymerases (Artsimovitch et al., 2011). The last group consists of compounds involved in the suppression of host defense responses. One such example is coronatine, a polyketide derivative mimicking the plant hormone jasmonic acid (JA) (Bender et al., 1999). JA is involved in several signaling pathways, one of which leads to the suppression of the salicylic acid (SA)-dependent defense pathway (Derksen et al., 2013). Another member of this group is the macrolactam syringolin A, which was demonstrated to be a proteasome inhibitor (Groll et al., 2008). Proteasome-mediated protein degradation is an essential part in a number of hormone-based signaling pathways in plants, including SA-mediated defense signaling (Kelley and Estelle, 2012).

Syringolin A and its minor variants are the products of a mixed non-ribosomal peptide synthetase (NRPS)/polyketide synthetase (PKS) encoded by the *sylA-E* gene cluster of certain *P. syringae* strains (Amrein et al., 2004; Ramel et al., 2009). It was first discovered in *P. syringae* pv. *syringae* (*Psy*) B301D-R (Amrein et al., 2004) and has so far been exclusively found among phylogroup II strains of this species (Baltrus et al., 2011). The compound contains a twelve-membered macrolactam ring containing a reactive double bond. The double bond interacts with active site residues of the eukaryotic proteasome, which results in covalent bond formation and inhibition of all three of its catalytic activities (Groll et al., 2008). Syringolin A was shown to suppress SA-mediated defense signaling and to counteract stomatal immunity in bean (*Phaseolus vulgaris*) and *Arabidopsis* (Schellenberg et al., 2010). Disruption of syringolin production was demonstrated to result in reduced lesion formation on bean (Groll et al., 2008) and in diminished wound entry and lesion spreading on *Nicotiana benthamiana* (Misas-Villamil et al., 2013).

In order to fully understand the mechanisms through which syringolin acts on plant cells and to determine its other potential roles in pathogenesis of *P. syringae*, we were aiming at establishing an infection model for a strain producing this compound in the well-studied model plant *Arabidopsis*, for which there exists a number of various genetic and "omics"-based tools. However, all our attempts to stably transform the *sylA-E* gene cluster into *P. syringae* pv. *tomato* (*Pto*) DC3000, a strain pathogenic to *Arabidopsis*, were not successful. Therefore, we decided to explore a different pathosystem involving one of the most important crop plants, common wheat (*Triticum aestivum*), for which we had available several independent isolates naturally containing the syringolin biosynthesis gene cluster (*Psy* B64 (Dudnik and Dudler, 2013a), *P. syringae* BRIP34876, and *P. syringae* BRIP34881 (Gardiner et al., 2013)), as well as a strain naturally lacking the phytotoxin:

(*Psy* SM (Dudnik and Dudler, 2013b)). It should be noted that both BRIP34876 and BRIP34881 were originally isolated from barley, but are as well able to cause disease on wheat (Gardiner et al., 2013).

The genomes of the four above-mentioned strains have been sequenced (Dudnik and Dudler, 2013a, 2013b; Gardiner et al., 2013), and based on multilocus sequence typing (MLST) analysis, they belong to phylogenetic clade II (Dudnik and Dudler, 2014). Genome analysis has also revealed that all four strains encode a relatively small complement of known T3Es: eleven effectors in BRIP34876 and BRIP34881, ten in *Psy* B64, and seven in *Psy* SM, as compared to 39 effectors in *P. syringae* pv. *tomato* DC3000 (Baltrus et al., 2011; Dudnik and Dudler, 2014). It is noteworthy that bacterial diseases of wheat are of low commercial significance and contribute only to about 3% of the total yield loss (Oerke and Dehne, 2004). However, having strains with a small set of T3Es provides an opportunity for determining roles of individual effectors due to presumably reduced functional overlap, as well as for pinpointing the most important defense pathways that need to be suppressed in order to allow disease progression. The small number of T3Es also suggests that these bacteria might be relying on non-type III secreted molecules to a greater extent than other *P. syringae* strains.

Therefore, this work focuses on characterization of virulence determinants in *P. syringae* strains with a small repertoire of type III effectors in wheat. Apart from the T3SS and effects of syringolin production, other important virulence-associated features were evaluated, such as the type VI secretion system and exopolysaccharide production.

6.3. Materials and methods

Bacterial strains and growth conditions. Bacterial strains and plasmids used in this study are listed in Table S1. Unless otherwise stated, bacteria were grown in Luria-Bertani (LB) medium (tryptone 10 g/liter, yeast extract 5 g/liter, and NaCl 5 g/liter) containing the appropriate amount of antibiotics: ampicillin (50 μg/mL), chloramphenicol (25 μg/mL), gentamicin (50 μg/mL), rifampicin (50 μg/mL), or tetracycline (15 μg/mL) at 28°C (*P. syringae*) or 37°C (*E. coli*) under constant agitation (220 rpm). For solid media 1.5% agar was used.

DNA manipulation. If not stated otherwise, standard procedures were used (Ausubel et al., 1998; Sambrook and Russell, 2001). Primers used in this work are listed in Supplementary Table S1. Chemicals and reagents were obtained from Sigma-Aldrich Chemie GmbH (Buchs, Switzerland).

All restriction enzymes and T4 DNA polymerase were purchased from New England Biolabs (Ipswich, MA, USA). Routine DNA amplification was carried out using Taq polymerase; PCR product and restriction digest product purification were performed using the GenElute™ PCR Clean-Up Kit; plasmid isolation was done using the GenElute™ HP Plasmid Purification Kit (all from Sigma-Aldrich Chemie GmbH, Buchs, Switzerland). Amplification of DNA fragments for the generation of mutagenic constructs was done using Extensor Long PCR Master Mix with Buffer 1 (Fisher Scientific Switzerland, Reinach, Switzerland).

Generation of pJQ200KSΔP*lac*. In order to avoid a possible polar effect in knock-out mutants due to the *lac* promoter of pJQ200KS, the corresponding DNA fragment was removed by digesting the vector with SphI and ApaI. The resulting fragments were separated by agarose gel electrophoresis, and the larger DNA fragment was cut out and purified using the Wizard® SV Gel and PCR Clean-Up System (Promega, Madison, WI, USA). Blunt ends were generated using T4 DNA polymerase in the presence of excess dNTPs. After heat inactivation, T4 DNA ligase (Promega, Madison, WI, USA) was added and the plasmid was re-ligated overnight at 4^{o}C. The ligation product was transformed into *E. coli* SURE electrocompetent cells (Agilent Technologies, Basel, Switzerland). The obtained plasmid has the ApaI site reconstituted and was verified by restriction digestion.

Generation of gene knock-out mutants by plasmid insertion. An approximately 700-bp-long fragment of a target gene was amplified using the respective P1 and P2 primers (Table S2). The primers were designed such that the inserted construct would disrupt a predicted conserved domain, where available. The fragment was then cloned into one of the suicide vectors, pJQ200KSΔP*lac* or pME3087, using respective restriction enzymes, and transformed into *E. coli* XL-1 Blue. After verification by sequencing, the construct was mobilized into *P. syringae* by tri-parental mating using *E. coli* HB101 (pRK600) as a helper strain. In order to generate the *sylC*_KO mutant of BRIP34876, the respective construct was first transformed into *E. coli* ST18, which then served as a donor in bi-parental mating. The *sylC*_KO mutant of *Psy* B64 was generated using bi-parental mating with *E. coli* S17-1 (pME3087-PS3) (Amrein et al., 2004). The location of the integrated construct was verified by PCR using the respective ch1 primer and either pJQ200KS_B1_R when pJQ200KSΔP*lac* was used, or pr_3087 H3 when pME3087 was used (See Table S2 for a complete list of primers and Table S3 for NCBI Genbank references).

Generation of markerless in-frame gene deletion mutants. Two approximately 700 bp-long fragments originating from upstream and downstream regions, respectively, of target gene's open

reading frame (ORF) were amplified using the respective primers (Table S2; primers P1 and P2 amplify the corresponding upstream fragments, primers P3 and P4 amplify the corresponding downstream fragments). The upstream and downstream fragments were joined by overlap extension PCR using primers P1 and P4, cloned into the suicide vectors pJQ200KSΔP*lac* using respective restriction enzymes, and transformed into *E. coli* XL-1 Blue. After verification by sequencing, the construct was mobilized into *P. syringae* by tri-parental mating using *E. coli* HB101 (pRK600) as a helper strain. Single recombinants of *P. syringae*, which were selected on plates containing the suitable antibiotic, were then grown on LB plates supplemented with 10% sucrose in order to select for double recombinants. Deletion of the target sequence was verified by PCR using the respective oriA and oriB primers, followed by sequence determination. See Table S2 for a complete list of primers and Table S3 for NCBI Genbank references. The resulting gene deletions retained an open reading frame (ORF) of eighteen to thirty bases, including the start and the stop codons.

Plant growth and infection assays. Common wheat (*Triticum aestivum*) cultivar Chinese Spring was grown under 16 hours light/8 hours dark conditions at 20^{o}C during the light phase and 18^{o}C during the dark phase without controlled humidity. For infiltration of 10-11-day-old plants, bacteria were grown overnight in 5 mL of LB medium. The culture was diluted the next day 1:25 in a total volume of 5 mL, grown further until an OD_{600} of 0.4-0.8 was reached, centrifuged at 2500x g, and re-suspended in sterile dH_2O. The suspension was diluted to 10^6 colony-forming units (CFU) per mL (OD_{600} 0.002) and infiltrated into primary leaves using a 3 mL needleless syringe. For surface (dip) inoculation, the overnight culture was diluted 1:25 in a total volume of 100 mL. The culture was further grown until an OD_{600} of 0.4-0.8 was reached, at which point it was harvested by centrifugation at 2500x g. The cells were washed once, and then re-suspended in dH_2O. The suspension was then diluted to 10^8 CFU/mL (OD_{600} 0.2), and 0.03% of the surfactant Silwet L-77 (Leu+Gygax AG, Birmenstorf, Switzerland) was added. Inoculation was performed by dipping primary leaves into the suspension for approximately 30 seconds.

Independently of the inoculation method used, counting of endophytic bacteria was performed as follows: a 3 cm-long leaf segment was cut out of the infiltrated area, surface-sterilized in EtOH for 15 seconds, and washed in dH_2O. Three leaf fragments infiltrated with the same strain were pooled together and macerated in 200 µL of sterile dH_2O using a sterilized pestle. Once only the major veins remained visible, 800 µL of dH_2O were added, and a 10-fold dilution series was generated. 50 µL of each dilution were plated onto LB-agar and incubated either for 42-44 hours at room

temperature (*Psy* SM and *Psy* B64), or for 24-26 hours at 28°C (BRIP34876). Each tested strain was assayed at least three times independently.

EPS, protease, and phytotoxin production assays. Alginate production was assayed by growing bacteria on Mannitol Glutamate Yeast (MGY) agar supplemented with 0.6 M sorbitol (Penaloza-Vazquez et al., 1997) at 28°C for 72 hours. Levan production was checked by incubating bacteria on MGY agar supplemented with 5% sucrose (Li and Ullrich, 2001) for seven days at 20°C. Protease production was measured by growing bacteria on 5% skim milk-agar plates for five days at 20°C. Syringolin production was assayed using the rice-based detection system as described by Ramel and colleagues (Ramel et al., 2009).

Bacterial competition assays. *P. syringae* and *E. coli* SURE (pJQ200KSΔP*lac*) were grown overnight in 5 mL of LB medium. The cultures were diluted the following day 1:25 for *Psy* and 1:100 for *E. coli* in a total volume of 5 mL and grown further until an OD_{600} of 0.4-0.8 was reached. Afterwards, cells were harvested and re-suspended in sterile dH_2O. The suspensions were diluted to an OD_{600} of 0.5 and each *P. syringae* strain was mixed with *E. coli* in a 1:1 ratio. The mixtures were spotted onto LB-agar without antibiotics and incubated at 28°C for 20 hours. Afterwards, the cells were scraped off, re-suspended in sterile dH_2O, and a dilution series was generated. 50 µL of each dilution were plated onto LB-agar supplemented with gentamycin and incubated overnight at 37°C in order to select for *E. coli*, followed by CFU counting.

6.4. Results

6.4.1. Syringolin production shows strain-specific effects on symptom development in wheat

In order to search for a potential role of syringolin A in the infection process on wheat, wild type (wt) *Psy* SM, *Psy* B64, and BRIP34876 were assayed. The latter two strains contain the *sylA-E* gene cluster (PssB64_04155 – PssB64_04151 and A979_11049 – A979_11069 respectively), while the first one does not contain it in its genome. These strains were compared with their following derivatives: *Psy* SM (pPL3syl) which contains the *sylA-E* gene cluster on a cosmid and is known to produce syringolin (Ramel et al., 2009), and *Psy* B64 *sylC*_KO and BRIP34876 *sylC*_KO, which are *sylC* knock-out mutants that are deficient in syringolin A production. Based on genome sequence comparison, strains BRIP34876 and BRIP34881 are nearly identical (Dudnik and Dudler, 2014), and therefore, only BRIP34876 was used. In addition, since surface inoculation is a more natural way of infection, as well as because syringolin A was demonstrated to be involved in counteracting

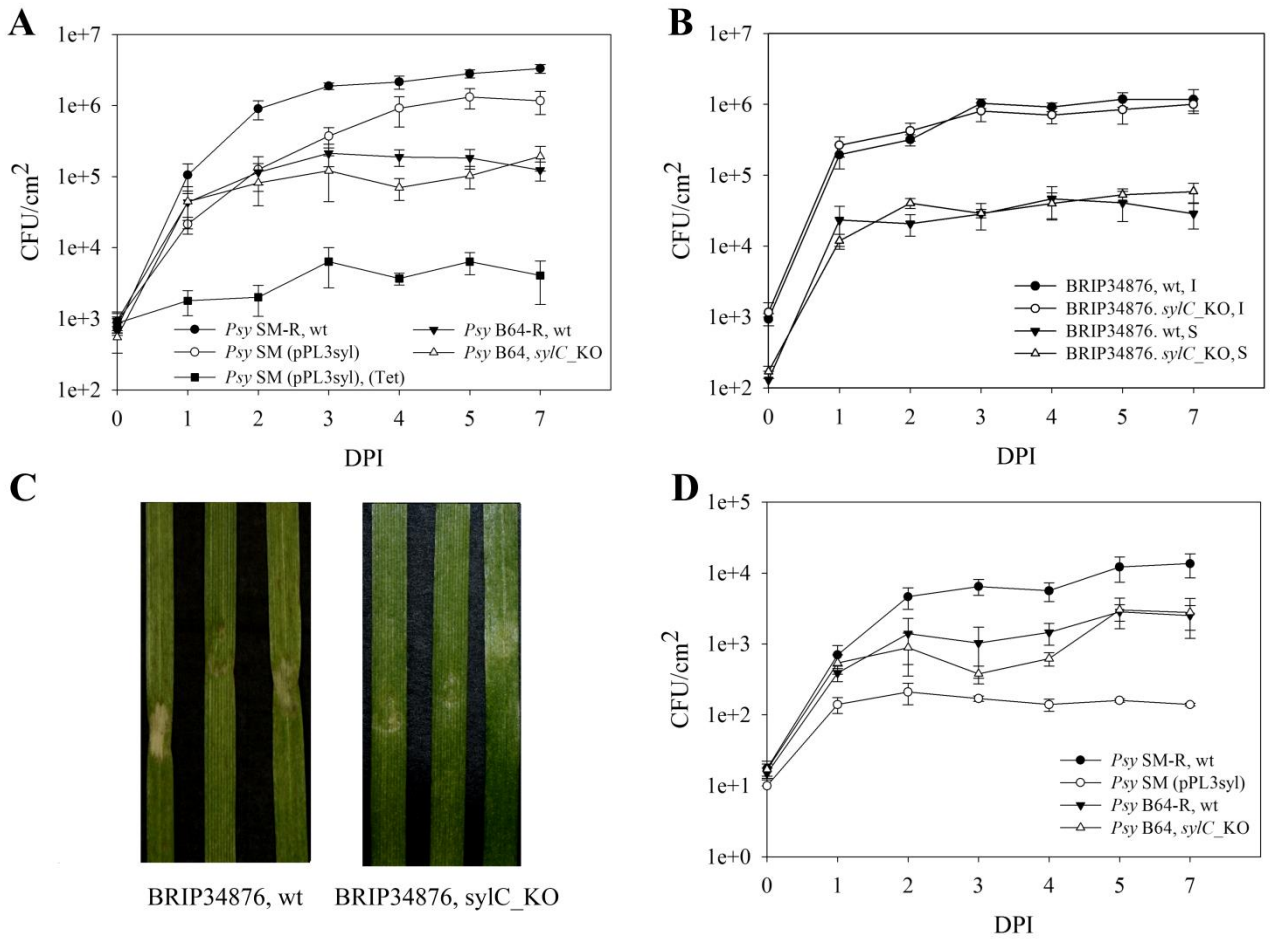


Figure 1. Endophytic bacterial population growth dynamics of *Psy* SM wt, *Psy* SM (pPL3syl), *Psy* B64 wt and *Psy* B64 *sylC*_KO after infiltration (**A**) and after surface inoculation (**D**). "wt" indicates wild type strains and "KO" indicates gene knock-out mutants obtained by plasmid insertion. "Tet" indicates that selection on tetracyclin-containing plates was performed. (**B**) Intra-foliar bacterial population dynamics of BRIP34876 wt and BRIP34876 *sylC*_KO after infiltration (I) and surface inoculation (S). CFU – colony forming units, DPI – day post inoculation (**C**) Disease symptoms caused by BRIP34876 wt (left) and BRIP34876 *sylC*_KO (right).

stomatal immunity in dicot species (Schellenberg et al., 2010), the strains were analyzed using both infiltration and surface inoculation. The results are presented in Figure 1.

When infiltrated directly into the leaf mesophyll, *Psy* B64 and its *sylC*-negative derivative did not show significant growth differences (Figure 1A). Infiltrated leaves remained green and completely asymptomatic with both strains (Figure S1A and B). Similarly, wild type BRIP34876 and the corresponding *sylC_KO*mutant also showed equal endophytic growth under these conditions (Figure 1B). However, in contrast to *Psy* B64, BRIP34876-infiltrated leaves exhibited more tissue damage and slightly higher level of chlorosis, which was less pronounced in the *sylC*-deficient mutant

(Figure 1C). In contrast, *Psy* SM (pPL3syl) showed a reduction in growth when compared to the wild type *Psy* SM, in particular during the early stages of infection (Figure 1A). Moreover, when the recovered bacteria were grown on LB-agar with tetracycline, which selects for bacteria carrying the cosmid pPL3syl, the strain showed a large decrease in cell numbers as compared to when the same suspensions were plated onto LB-agar alone (Figure 1A). This suggests loss of pPL3syl under non-selective conditions, which was further verified by PCR using *sylC*-specific primers (data not shown). This finding suggests that there is a strong negative selection against *Psy* SM cells carrying pPL3syl upon infiltration into wheat leaves. Infection with *Psy* SM caused small lesions spread throughout the entire infiltrated area, without a pronounced chlorosis. Leaves infiltrated with *Psy* SM (pPL3syl) showed a slightly smaller number of lesions (Figure S1C and 1D, respectively).

When bacteria were inoculated by dipping, no difference in endophytic growth was detected between the wild type and the *sylC* knock-out mutants of both BRIP34976 and *Psy* B64 and (Figure 1B and 1D respectively). Interestingly, *Psy* SM (pPL3syl) showed an overall reduction of bacterial counts, even without applying the tetracycline selection, which suggests that syringolin production in this strain might prevent it even from entering the mesophyll. In addition, none of the strains was able to cause visible symptoms (data not shown). As obvious from Figure 1, after surface inoculation, none of the strains was able to grow to the same cell densities as observed upon direct infiltration.

In conclusion, the influence of syringolin production on leaf colonization appears to be strain-specific, and while in the case of BRIP34876 syringolin A production seemed to enhance symptom development, it had no or even a detrimental effect on the endophytic growth of *Psy* B64 and *Psy* SM, respectively (Figure 1). This is, however, only seen when the strains are infiltrated directly into leaves, as bacterial densities after surface inoculation did not reach high-enough values to induce visible symptoms.

6.4.2. The T3SS is the key pathogenicity determinant

A relatively small number of encoded T3Es is a property described for phylogroup II strains (Baltrus et al., 2011). However, the wheat and barley isolates appear to have a particularly minimized repertoire of the T3Es: eleven in BRIP34876 and BRIP34881, ten in *Psy* B64, and seven in *Psy* SM (Dudnik and Dudler, 2014). Therefore, it is possible that these strains rely to a larger extent on non-type III secreted molecules, such as phytotoxins or type V/type VI effectors. In order to test this

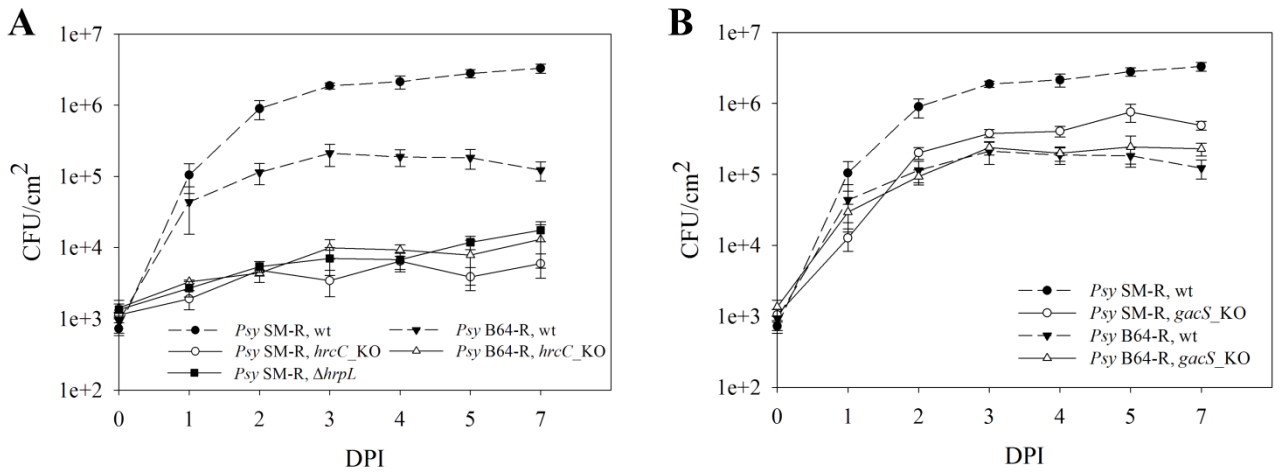


Figure 2. Endophytic bacterial population growth dynamics of (**A**) T3SS-deficient (*hrcC*_KO or Δ*hrpL*) *Psy* SM and *Psy* B64, and (**B**) *gacS*-deficient *Psy* SM and *Psy* B64 (*gacS*_KO). Data for the corresponding wt strains is given as a reference (dashed lines).

hypothesis, T3SS-deficient mutants were generated by knocking out the *hrcC* gene, which encodes an essential structural component of the T3SS (Deng and Huang, 1999). *Psy* B64 and *Psy* SM were chosen for this analysis because the two strains had the smallest number of T3Es among the strains available to us. Interestingly, knocking out the T3SS in both strains nearly abolished their ability to reproduce *in planta*, thus rendering them non-pathogenic (Figure 2A). The same result was observed upon deletion of the *hrpL* gene in *Psy* SM, which encodes the major transcriptional regulator of the T3SS (Tang et al., 2006). This shows that the type III secretion system still remains essential for the infection progress, and the small effector repertoire is sufficient to allow mesophyll colonization.

Regulation of the T3SS in *P. syringae* differs between strains, and in the model strain *P. syringae* pv. *tomato* (*Pto*) DC3000, it is under control of the GacA/GacS two-component system (Chatterjee et al., 2003). This two-component system is a global regulator in *P. syringae* and other pseudomonads (Lapouge et al., 2008) controlling, among others, quorum sensing, motility, phytotoxin and protease production (Chatterjee et al., 2003; Kitten et al., 1998; Ramel et al., 2012). In contrast, in *P. syringae* pv. *syringae* B728a, T3SS regulation appears to be uncoupled from the regulation by GacS (Records and Gross, 2010; Willis et al., 1990). Both *Psy* SM and *Psy* B64 are more related to *Psy* B728a than to *Pto* DC3000 (Dudnik and Dudler, 2014), and therefore it is likely that the wheat isolates also have the T3SS functioning independently of GacA/GacS regulation. Such setting would present a good opportunity to study an influence of non-type III secreted virulence factors, and therefore knock-out mutants of the *gacS* gene were constructed in the two strains.

When tested for endophytic growth, *Psy* SM *gacS*_KO showed an approximately eight-fold reduction in endophytic growth after infiltration (Figure 2B). In addition, *Psy* SM *gacS*_KO infiltrations resulted in a minor reduction of lesion numbers (Figure S1E). Surprisingly, knocking out the *gacS* gene in *Psy* B64 had no effect at all on mesophyll colonization of this strain (Figure 2B). Spontaneous inactivation of the GacA/GacS two-component system is a well-known phenomenon among the pseudomonads (van den Broek et al., 2005; Driscoll et al., 2011), and therefore could be a potential reason for the observed phenotype. To further investigate this, *Psy* B64 was tested for syringolin A production using our rice assay (*pir7b* expression induction in rice in the presence of syringolins (Ramel et al., 2009)), as well as for protease production by growing the strain on milk-agar. Both tests yielded negative results (Figure S2), thus supporting the hypothesis that the GacA/GacS system in *Psy* B64 is not active, in spite of the fact that, based on our genome sequence analysis, the respective genes appear to be intact. The lack of a functional GacA/GacS two-component system would explain the observed differences in overall endophytic growth of the wild type *Psy* SM and *Psy* B64 (Figure 1A). In addition, this would also explain the lack of any visible lesions on *Psy* B64-infiltrated leaves, as *gacA* or *gacS*-negative mutants in *P. syringae* are known to show a significant reduction in disease symptom intensity (Chatterjee et al., 2003; Willis et al., 1990).

6.4.3. Importance of individual T3Es

Because of the small effector repertoire and ease of genetic manipulation we decided to further characterize *Psy* SM. The genome of this strain encodes the following seven known T3Es: AvrE1, HopM1, HopI1, HopAA1, HopBA1, HopAZ1, and HopA2 (Dudnik and Dudler, 2013b). Only two of these are well-characterized: HopM1 targets ADP ribosylation factor-guanine nucleotide exchange factor (ARF-GEF) proteins for degradation (Gangadharan et al., 2013; Nomura et al., 2006), and HopI1 is involved in the degradation of the Hsp70 chaperon (Jelenska et al., 2010). Out of the remaining ones, AvrE1 is hypothesized to mimic activated G-proteins (Ham et al., 2009), while HopAA1 is a necrosis-inducing effector with a putative GTPase-activating protein (GAP) domain (Munkvold et al., 2009) that was demonstrated to enhance epiphytic growth and survival in tomato (Lee et al., 2012b). However, no exact molecular mechanism has been described so far for the latter two. These four effectors are found in almost all of the currently sequenced *P. syringae* strains (Baltrus et al., 2011), and although disrupted or truncated in some strains, they are likely involved in the modulation of conserved host defense pathways. Moreover, mutants of these four

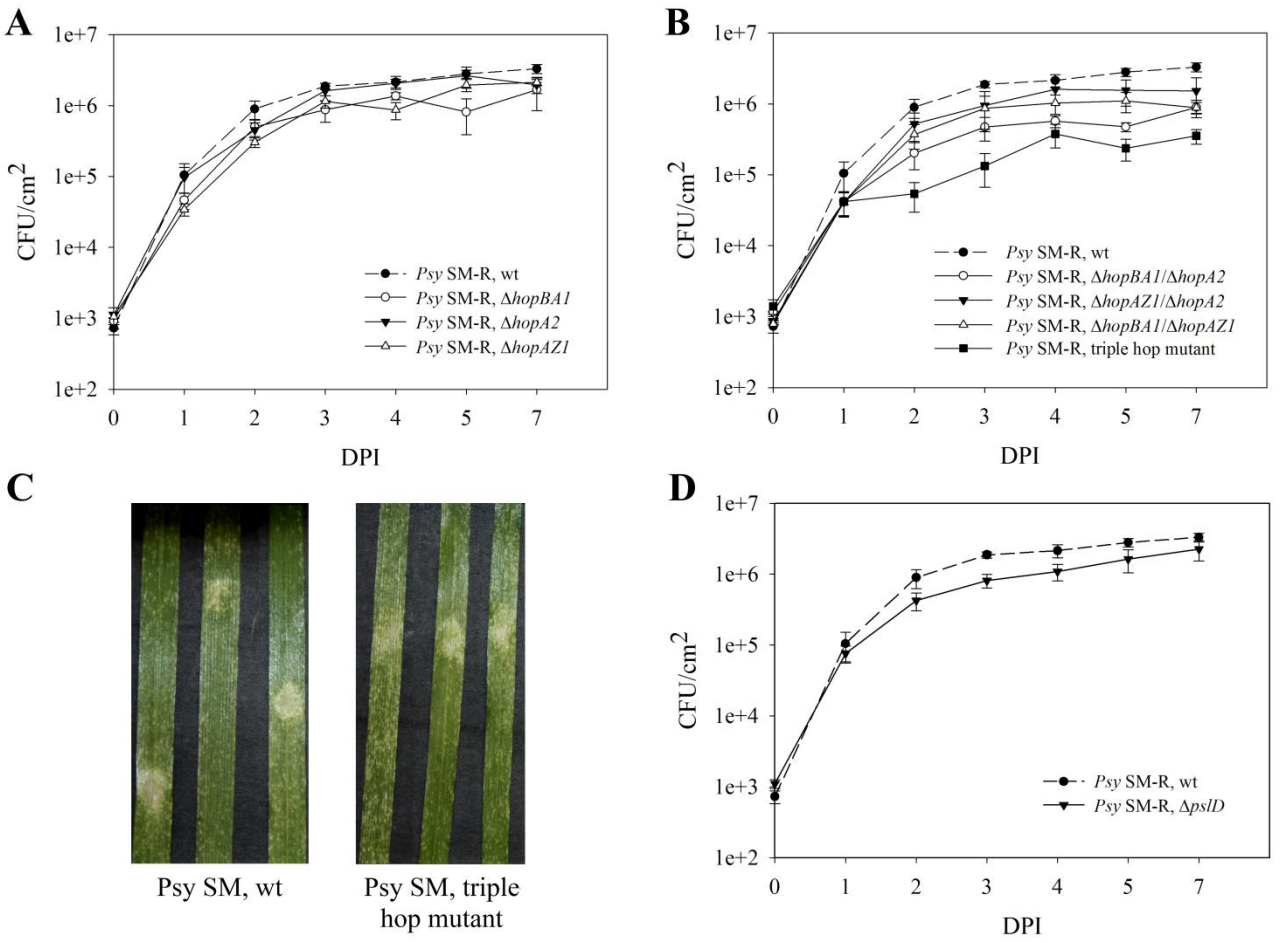


Figure 3. Endophytic bacterial population growth dynamics of (**A**) individual T3E mutants, (**B**) double and the triple T3E mutants, and (**D**) Δ*pslD* mutant. Data for *Psy* SM wt is given as a reference (dashed lines). (**C**) Disease symptoms caused by *Psy* SM wt (left) and its triple hop mutant (right).

"core" effectors have been characterized in terms of the impact on growth and disease progression (Badel et al., 2006; Jelenska et al., 2007). Therefore, we decided to focus on the three remaining effectors: HopA2, HopAZ1, and HopBA1. In order to investigate the importance of the selected effectors, individual markerless gene deletion mutants were constructed. Furthermore, all possible combinations of double mutants and the triple mutant were generated as well.

When infiltrated into wheat leaves, the individual Hop mutants resulted in a small endophytic growth reduction when compared to the wild type (Figure 3A). The strongest impact was observed when *hopBA1* was deleted, which caused an approximately two-fold growth reduction, while the Δ*hopA2* strain was nearly indistinguishable from the wild type. The double mutants showed a further reduction of endophytic growth, which was most pronounced in the Δ*hopBA1*Δ*hopA2* strain, where bacterial counts were four to five times smaller as compared to the wild type (Figure 3B). The

triple mutant showed an *in planta* growth reduction of more than 10-fold (Figure 3B). Interestingly, this became apparent only two days after inoculation, which might imply that these effectors become crucial only when the bacterial population reaches a certain size, to allow further expansion. What is also notable is that leaves infiltrated with the triple mutant showed a degree of lesion formation similar to the wild type *Psy* SM-R, despite the reduced bacterial density (Figure 3C).

Contribution of exopolysaccharides to virulence of Psy SM on wheat

P. syringae was demonstrated to produce two types of exopolysaccharides (EPS): levan and alginate (Osman et al., 1986). Levan is synthesized from sucrose by a single secreted enzyme termed levansucrase (Li and Ullrich, 2001). The genome of *Psy* SM contains two levansucrase-encoding genes, *lscA* and *lscC*. Alginate biosynthesis involves several enzymes (Franklin et al., 2011), the majority of which are found within a gene cluster (Penaloza-Vazquez et al., 1997). In addition, like the majority of sequenced *P. syringae* strains, the genome of *Psy* SM contains homologs of *psl* genes from *P. aeruginosa*. Psl is a complex polysaccharide consisting repeating blocks of D-glucose, D-mannose, and L-rhamnose (Byrd et al., 2009) that was demonstrated to be involved in biofilm formation of *P. aeruginosa* (Jackson et al., 2004). EPS are presumed to be involved in the protection from desiccation, epiphytic fitness, and osmotolerance, all of which were demonstrated for alginate (Penaloza-Vazquez et al., 1997; Schnider-Keel et al., 2001; Yu et al., 1999). Levan was shown to be required for lesion formation by *Erwinia amylovora* on pear seadlings (Geier and Geider, 1993), however, in *P. syringae* it is rather presumed to be a storage polymer (Laue et al., 2006). The *psl* gene cluster was identified to be transcriptionally active in *P. syringae* (Hockett et al., 2013; Records and Gross, 2010), however, no further characterization was performed.

When assayed on MGY-agar with 0.6 M sorbitol, a medium which stimulates alginate production (Penaloza-Vazquez et al., 1997), *Psy* SM appeared to be non-mucoid (Figure S3A). This was unexpected as *Psy* SM contains a full complement of *alg*/*muc* genes which are 99-100% identical to their corresponding homologs from the alginate-producing strain BRIP34876. Since no alginate production was detected in wild type *Psy* SM, the Δ*algA* strain was excluded from the study. Apart from that, *Psy* SM is levan-positive, and in the *lscC* deletion mutant we constructed, levan production was completely abolished (Figure S3B). No effect of *lscA* deletion on levan production was observed on MGY+5% sucrose plates at 20^{o}C, which is the temperature used for cultivating test plants. This is in line with the previously published data (Li and Ullrich, 2001). However, Hockett and colleagues (Hockett et al., 2013) reported that *lscA* was transcriptionally active in *Psy* B728a,

and thus, we also constructed a double Δ*lscA*/Δ*lscC* mutant, which had the same phenotype as the single Δ*lscC* mutant (Figure S3B). In order to test whether Psl production is required for virulence of *Psy* SM, we generated a deletion mutant of the *pslD* gene, which was demonstrated to be essential for Psl production in *P. aeruginosa* (Byrd et al., 2009; Friedman and Kolter, 2004). Unfortunately, there is no proper assay for Psl production available for *P. syringae*, as regulation is potentially different from that of *P. aeruginosa*.

To investigate the importance of the EPS in wheat mesophyll colonization, the Δ*pslD* and Δ*lscA*/Δ*lscC* mutants were assayed using syringe-infiltration. The levansucrase-deficient mutant showed a negligible reduction in intra-foliar bacterial counts only at around 3-4 DPI (Figure S4). The Δ*pslD* strain showed a minor, upproximately two-fold endophytic growth reduction throughout the entire time monitored (Figure 3D).

6.4.4. The type VI secretion system is involved in inter-species competition, but not in virulence

The type VI secretion system (T6SS) is a multipurpose protein delivery machinery frequently found in genomes of pathogenic and non-pathogenic Gram-negative bacteria (Kapitein and Mogk, 2013). In several pathogens, the T6SS was shown to be involved in virulence to a variable extent (Burtnick et al., 2011; Mulder et al., 2012; Rao et al., 2004). However, it was additionally demonstrated to be involved in inter-species competition, biofilm formation, "self" versus "non-self" discrimination, and protection from predation (Coulthurst, 2013; Kapitein and Mogk, 2013). In *Pto* DC3000, the T6SS was shown to have no function in virulence, but rather to be involved in competition with other microbial species, including yeast (Haapalainen et al., 2012).

The genomes of both *Psy* SM and *Psy* B64 contain two independent T6SS gene clusters (Dudnik and Dudler, 2014). Our analysis indicates that due to a large inversion in the *Psy* B64 genome, the T6SS-1_{SM} is homologous to the T6SS-2_{B64}, and vice versa. Based on sequences of the respective VipA (TssA) proteins, the T6SS-1 of *Psy* SM appears to be related to the HSI-II of *P. aeruginosa*, whereas the T6SS-2 shows homology to the *evpA–H* gene cluster from *Edwardsiella tarda*. It should be noted that it is not uncommon that T6SS gene clusters are only active under certain conditions(Bernard et al., 2010; Burtnick et al., 2011; Haapalainen et al., 2012). In order to investigate whether any of the two gene clusters was active, single and double T6SS knock-out

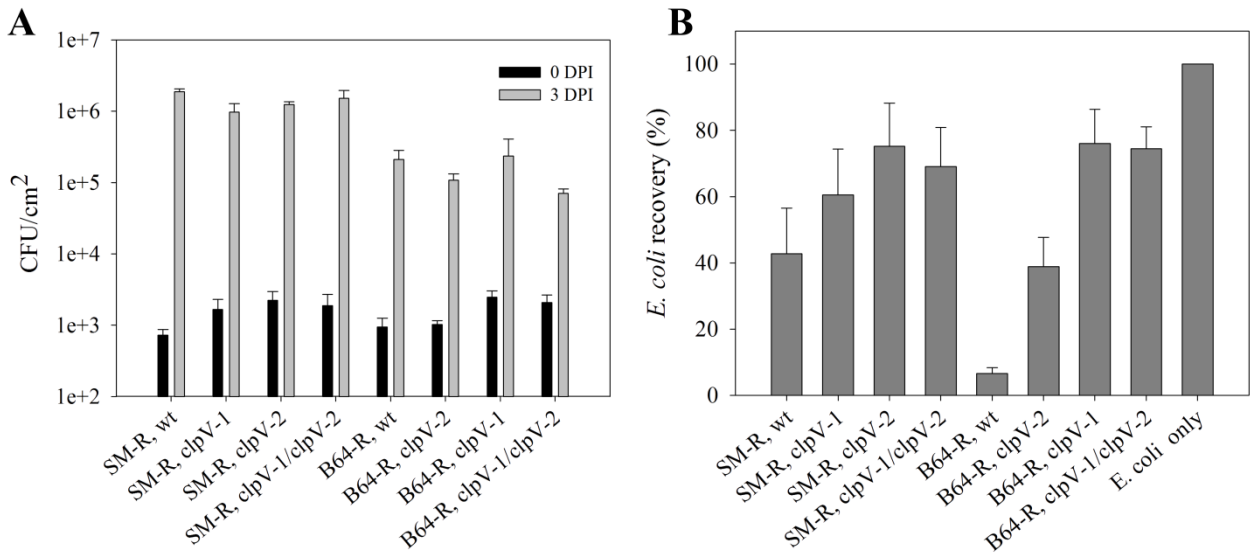


FIGURE 4. (**A**) Endophytic population sizes of *Psy* SM wt, *Psy* B64 wt, and the corresponding T6SS-deficient mutants at 3 DPI. (**B**) Inhibition of *E. coli* growth seen after co-incubation with *Psy* SM wt, *Psy* B64 wt, and the corresponding T6SS-deficient mutants.

mutants were constructed in *Psy* SM and *Psy* B64. The mutagenic constructs were targeting the *clpV* ATPases, which are known to be essential for T6SS function (Hood et al., 2010; Records and Gross, 2010).

When tested *in planta* by infiltration, the mutants behaved like the wild type strains (Figure 4A), and no significant differences were observed in terms of endophytic growth. Since a similar effect was observed in *Pto* DC3000 (Haapalainen et al., 2012), this question was not addressed any further. To explore whether the T6SS had an influence on the ability of *Psy* SM and *Psy* B64 to compete with other bacteria, co-incubation experiments with *E. coli* were performed. To do so, overnight cultures were diluted, grown to mid-logarithmic phase and washed in dH_2O. After the optical density of each culture was adjusted to 0.5, the suspensions of *E. coli* and *Psy* were mixed in a 1:1 ratio and spotted on LB-agar plates. The plates were incubated at 28°C for 20 hours, after which bacteria were re-suspended in dH_2O. Dilution series were then prepared and plated on a medium selective for *E. coli*. The results of these experiments show that T6SS mutants had a significantly reduced ability to suppress *E. coli* multiplication, in particular for *Psy* B64 (Figure 4B). While co-incubating the wild type *Psy* B64 resulted in an approximately 16-fold reduction of *E. coli* recovery (as compared to *E. coli*-only cultures), co-incubation with the double *clpV*-1_{B64}/*clpV*-2_{B64} mutant resulted in about 75% recovery of *E. coli*. Interestingly, the *clpV*-1_{B64} single gene knock-out had the same effect as the double *clpV*$_{B64}$ mutant (p=0.72, t-Test), whereas the *clpV*-2_{B64} single mutant had a smaller effect

(Figure 4B). In contrast to *Psy* B64, the wild type *Psy* SM inhibited *E. coli* growth to a much lesser extent (43% recovery on average). Similarly to *Psy* B64, there was no cumulative effect of knocking out both $clpV_{SM}$ genes observed, and the double $T6SS_{SM}$ mutant showed the same degree of inhibition as the *clpV*-2_{SM} single mutant.

6.5. Discussion

While it is generally assumed that phytotoxin production in *P. syringae* is not absolutely required for the infection process, these compounds have been demonstrated to enhance disease progression and symptom development (Arrebola et al., 2009; Bender et al., 1999). Here, we have demonstrated that, when produced in wild type amounts, syringolin promotes lesion formation on wheat. This also confirms the previous findings by Groll and colleagues (Groll et al., 2008), who reported that lack of syringolin production resulted in reduced lesion formation on bean. Furthermore, we demonstrate that decreased lesion formation is not due to a reduction of the intra-foliar bacterial density. At present we can only speculate why this is the case. One possible explanation is that presence of syringolin allows a more efficient effector translocation and/or production of necrosis-inducing toxins such as syringomycin, which in turn would result in a higher level of damage to host tissue. It is also interesting is that over-production of syringolin by *Psy* SM (pPL3syl), which occurs due to presence of several copies of pPL3syl (Ramel et al., 2009), is detrimental for the infection process and restricts the growth of this strain soon after inoculation. A possible reason is that the bacterium actually requires some host proteasomes for the proper functioning of some of its T3Es. As mentioned previously, HopM1 is one such example (Nomura et al., 2011), and it is likely that in wheat, this effector is essential during the early stages of infection. The results obtained with the triple hop mutant (Δ*hopBA1*/Δ*hopA2*/Δ*hopAZ1*) support this statement, as the effect of the deletions only becomes apparent at 2 DPI. Therefore, it would be interesting to identify the targets of HopM1 in wheat as it might lead to discovery of novel early pathogen-response signaling pathways in this organism. The other three core T3Es (HopAA1, HopI1 and AvrE1), based on the currently available data, do not appear to depend on proteasome function (Ham et al., 2009; Jelenska et al., 2010; Munkvold et al., 2009).

Another interesting observation was the effect of the *hopBA1* deletion, which alone, or in combination with the other non-core T3Es, would result in a somewhat lower virulence when compared to other combinations. Moreover, *hopBA1* is common to all currently published phylogroup II *P. syringae* strains isolated from grasses (family Poaceae) (Dudnik and Dudler, 2014).

Thus, based on the obtained results, this effector indeed appears to play an important role in colonization of the wheat leaf mesophyll. Hence, further studies on this protein could provide some interesting insights into its exact role in this process, as well as on Poaceae-specific immunity mechanisms.

It is also apparent from our results that the seven identified T3Es are indeed sufficient for *Psy* SM to allow proper endophytic growth on wheat. Based on the growth of the *gacS*_KO mutant, it appears that the relative contribution of the GacA/GacS regulon to the overall virulence is about 30%, with a further approximately 30% being contributed by the three non-core T3Es, and the rest being potentially attributed to the four core T3Es. This, however, does not completely exclude the possibility that the strain possesses other, so far unknown, effectors, but if indeed present, their contribution to the overall virulence appears negligible. This finding is in good agreement with previously published work on the minimal effector repertoire required by *Pto* DC3000 to grow on *N. benthamiana* (Cunnac et al., 2011), where the authors determined that only eight T3Es out of about three dozen were sufficient to reach bacterial cell counts comparable to those of the wild type strain. One of the required T3Es was HopM1. This work also demonstrated that in some cases, addition of a single effector resulted only in a minor increase of endophytic growth. Here, a similar trend was observed, as the analyzed single T3E mutants only showed a minor decrease in intra-foliar bacterial growth. Nonetheless, even though the seven effectors allow *Psy* SM to reach a relatively high population density, it also became apparent that they were unable to fully suppress host immunity. The lesions, which were at first assumed to be disease symptoms, appeared in somewhat higher numbers in the triple hop mutant, even though lower population densities were reached. Hence, the reasonable conclusion is that upon reaching a certain cell density, even the wild type was not fully able to down-regulate the host immune response. The triple mutant is further impaired in this ability, thus, the lesions appear at even lower bacterial densities.

Another notable finding is that deletion of the *pslD* gene resulted in a decrease of endophytic growth, albeit in a minor one. It has been demonstrated previously that neither alginate nor levan are essential for biofilm formation in *P. syringae*, and that there must be another EPS that is involved in this process (Laue et al., 2006). The EPS Psl was discovered around that time (Friedman and Kolter, 2004; Jackson et al., 2004; Matsukawa and Greenberg, 2004) and a homologous biosynthesis gene cluster has been identified in *P. syringae* later on. The *psl* gene cluster in *P. syringae* is not exactly identical to that of *P. aeruginosa*: the *pslC* gene homolog is found at a different chromosomal

location, and another essential gene, *pslL*, is absent. Up to now there are only transcriptional data available for this gene cluster (Hockett et al., 2013; Records and Gross, 2010), and to our knowledge, this is the first attempt to demonstrate its role experimentally. It is possible that there are only minor amounts of this EPS produced, making it difficult to detect. This would also explain why it has not been characterized so far, and that its contribution to virulence appears to be a minor one. It is possible that with a small number of T3Es, even a minor change in virulence factor composition becomes apparent. Nevertheless, a more elaborate study is required to draw definite conclusions regarding the production and the role of Psl in virulence of *P. syringae*.

In this work we demonstrated that both *Psy* SM and *Psy* B64 have at least one active T6SS gene cluster that the strains could use against other bacteria. The results obtained suggest that the major role in this process is played by T6SS-2_{SM} (T6SS-1_{B64}), as knock-out mutants of it loose the ability to compete with *E. coli* to the same extent as the double mutants. Based on the homology of the respective TssA proteins, T6SS-1_{SM} (T6SS-2_{B64}) is more related to HSI-1 of *Pst* DC3000, which was as well demonstrated not to be involved in competition with other microbes (Haapalainen et al., 2012). The effects observed in the knock-out mutants of T6SS-1_{SM} (T6SS-2_{B64}) could also result from indirect or pleiotropic effects, as was observed for the T6SS in other bacteria (Lertpiriyapong et al., 2012; Sha et al., 2013). An alternative explanation is that the two T6SS share some of their translocated substrates, even though no translocon overlap between different T6SS has been demonstrated (Burtnick et al., 2011; Hachani et al., 2011). It is, however, possible that such an overlap may have been overlooked since at present, very little is known about type VI effectors. It is interesting that the two related strains show a large difference in the extent of their ability to inhibit *E. coli* multiplication. This is most likely due to differences in type VI effector compositions (Dudnik and Dudler, 2014). It is noteworthy that the inactivation of the T6SS did not completely suppress the cytotoxic activity towards *E. coli*. This is could either be due to a more efficient nutrient utilization by *Psy*, or due to the action of bacteriocins produced by these strains (Dudnik and Dudler, 2014).

6.6. Supplementary figures and tables

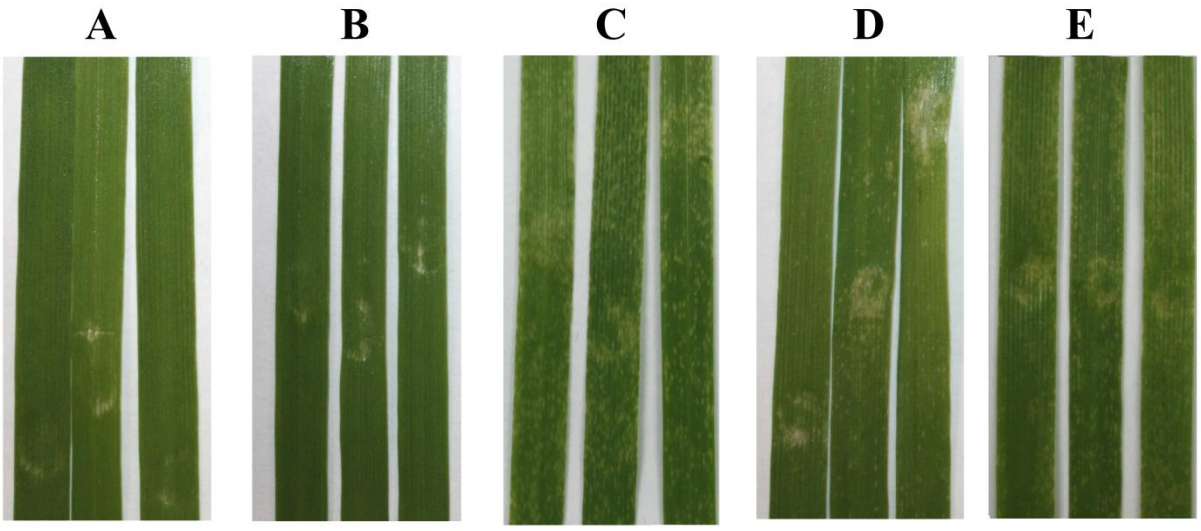


Figure S1. Wheat leaves after infiltration with (**A**) *Psy* B64 wt, (**B**) *Psy* B64 *sylC*_KO, (**C**) *Psy* SM wt, (**D**) *Psy* SM (pPL3syl), (**E**) *Psy* SM *gacS*_KO at 7 DPI

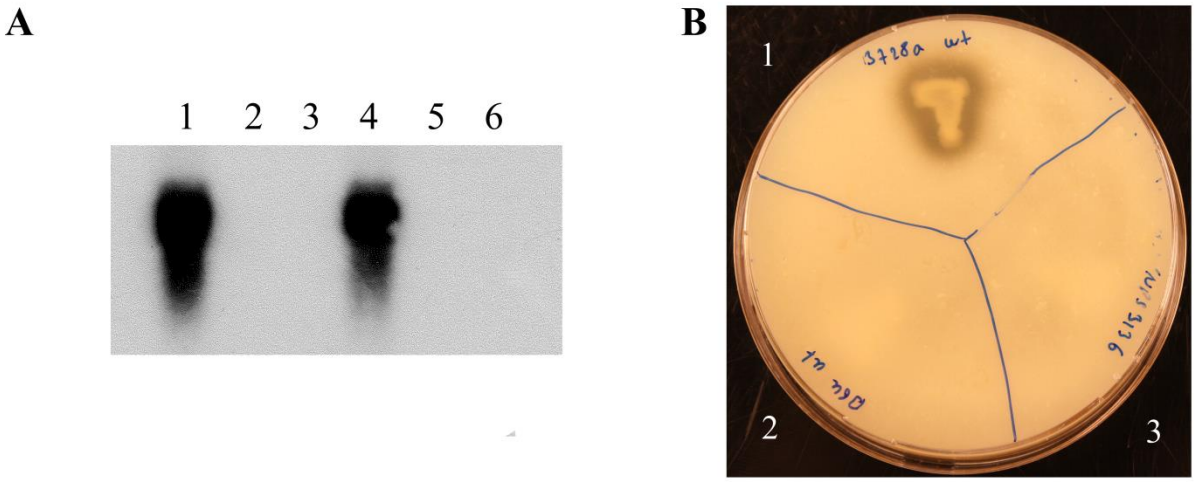


Figure S2. (**A**) Northern blot depicting expression levels of *pir7b* gene transcript in rice after 16h post- infiltration with *P. syringae*. This gene gets expressed only in the presence of syringolin A. (1) and (4) *Psy* B301D-R wt, (2) and (5) *Psy* B64 wt, (3) and (6) *Psy* B64 *gacS*_KO. (**B**) Milk-agar plates. The halo occurs due to protease secretion. (1) *Psy* B728a wt, (2) *Psy* B64 wt, (3) NPS3136 (*Psy* B728a *gacS*::Tn5)

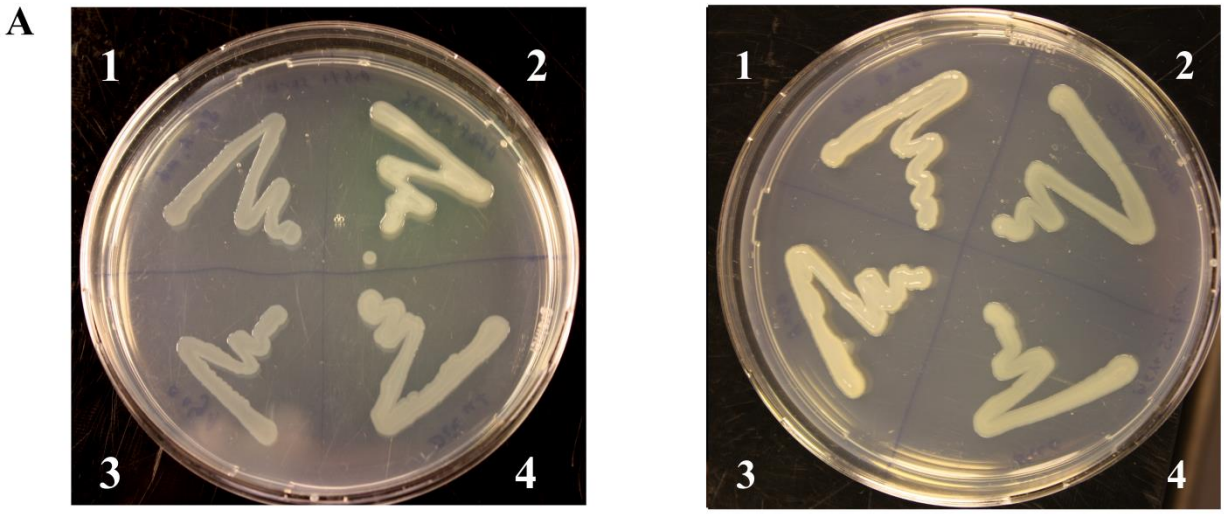


Figure S3. (**A**) Alginate production assay using MGY-agar plates supplemented with 0.6M sorbitol: (1) *Psy* SM wt, (2) BRIP34876 wt, (3) *Psy* SM Δ*algA*, (4) *Psy* B64 wt. (**B**) Levan production assay using MGY-agar plates supplemented with 5% sucrose: (1) *Psy* SM wt, (2) *Psy* SM Δ*lscA*/Δ*lscC*, (3) *Psy* SM Δ*lscA*, (4) *Psy* SM Δ*lscC*.

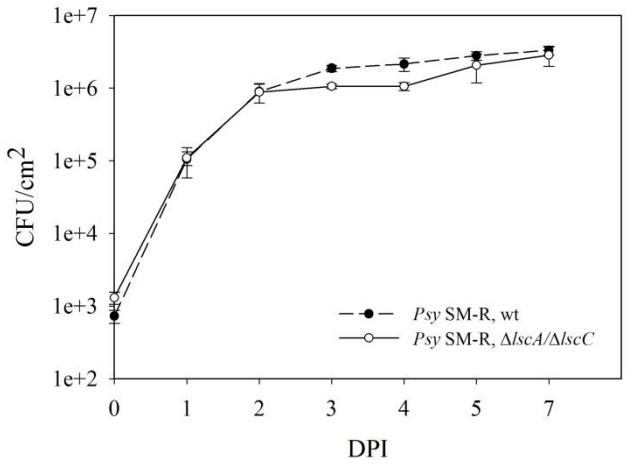


Figure S4. Endophytic bacterial population growth dynamics of *Psy* SM Δ*lscA*/Δ*lscC*. Data for *Psy* SM wt is given as reference (dashed lines).

Table S1. List of strains used in this study.

Strain or Plasmid	Genotype, relevant characteristics	Source or Reference
Pseudomonas syringae		
Psy SM	Pathovar *syringae*, wild type strain, isolated from wheat	(Dudnik and Dudler, 2013b)
Psy SM (pPL3syl)	*Psy* SM with *sylA-E* gene cluster on pLAFR3	(Ramel et al., 2009)
Psy SM-R	Spontaneous Rif^R mutant of *Psy* SM	This study
Psy SM-R, *gacS*_KO	*gacS* knock-out mutant of *Psy* SM-R, $Gent^R$	This study
Psy SM-R, *hrcC*_KO	*hrcC* knock-out mutant of *Psy* SM-R	This study
Psy SM-R, Δ*hrpL*	*hrpL*deletion mutant of *Psy* SM-R	This study
Psy SM-R, Δ*hopBA1*	*hopBA1*deletion mutant of *Psy* SM-R	This study
Psy SM-R, Δ*hopA2*	*hopA2* deletion mutant of *Psy* SM-R	This study
Psy SM-R, Δ*hopAZ1*	*hopAZ1*deletion mutant of *Psy* SM-R	This study
Psy SM-R, Δ*hopBA1*/Δ*hopA2*	*hopBA1*/*hopA2* deletion mutant of *Psy* SM-R	This study
Psy SM-R, Δ*hopBA1*/Δ*hopAZ1*	*hopBA1*/*hopAZ1* deletion mutant of *Psy* SM-R	This study
Psy SM-R, Δ*hopAZ1*/Δ*hopA2*	*hopAZ1*/*hopA2* double deletion mutant of *Psy* SM-R	This study
Psy SM-R, triple hop mutant	*hopBA1*/*hopA2*/*hopAZ1* triple deletion mutant of *Psy* SM-R	This study
Psy SM-R, Δ*algA*	*algA* deletion mutant of *Psy* SM-R	This study
Psy SM-R, Δ*lscA*/Δ*lscC*	*lscA*/*lscC* double deletion mutant of *Psy* SM-R	This study
Psy SM-R, Δ*pslD*	*pslD* deletion mutant of *Psy* SM-R	This study
Psy SM-R, *clpV*-1_KO	*clpV*-1 knock-out mutant of *Psy* SM-R, Tet^R	This study
Psy SM-R, *clpV*-2_KO	*clpV*-2 knock-out mutant of *Psy* SM-R, $Gent^R$	This study
Psy SM-R, *clpV*-1_KO/*clpV*-2_KO	*clpV*-1/*clpV*-2 double knock-out mutant of *Psy* SM-R, $Gent^R$/Tet^R	This study
Psy B64	Pathovar *syringae*, wild type strain, isolated from wheat	(Dudnik and Dudler, 2013a)
Psy B64, *sylC*_KO	*sylC* knock-out mutant of *Psy* B64, Tet^R	This study
Psy B64-R	Spontaneous Rif^R mutant of *Psy* B64	This study
Psy B64-R, *gacS*_KO	*gacS* knock-out mutant of *Psy* B64-R, $Gent^R$	This study
Psy B64-R, *hrcC*_KO	*hrcC* knock-out mutant of *Psy* B64-R, $Gent^R$	This study
Psy B64-R, *clpV*-1_KO	*clpV*-1 knock-out mutant of *Psy* B64-R, $Gent^R$	This study
Psy B64-R, *clpV*-2_KO	*clpV*-2 knock-out mutant of *Psy* B64-R, $Gent^R$	This study
Psy B64-R, *clpV*-1_KO/*clpV*-2_KO	*clpV*-1/*clpV*-2 double knock-out mutant of *Psy* B64-R, $Gent^R$/Tet^R	This study
BRIP34876	Wild type, isolated from barley	(Gardiner et al., 2013)
BRIP34876, *sylC*_KO	*sylC* knock-out mutant of BRIP34876, $Gent^R$	This study
Psy B301D-R	Pathovar *syringae*, wild type strain, isolated from pear, Rif^R	(Cody and Gross, 1987)
Psy B728a	Pathovar *syringae*, wild type strain, isolated from bean, Rif^R	(Feil et al., 2005)
NPS3136	*gacS* knock-out mutant of *Psy* B728a, Rif^R/Kan^R	(Willis et al., 1990)
Escherichia coli		
XL-1 Blue	*endA*1 *gyrA*96(nal^R) *thi*-1 *recA*1 *relA*1 *lac glnV*44 F'[::Tn10 $proAB^+$ $lacI^q$ Δ(*lacZ*)M15] *hsdR*17(r_K^- m_K^+)	(Bullock et al., 1987)

HB101 (pRK600)	F$^-$ *mcrB mrr hsdS*20(r_B^- m_B^-) *recA*13 *leuB*6 *ara*-14 *proA*2 *lacY*1 *galK*2 *xyl*-5 *mtl*-1 *rpsL*20(Sm^R) *glnV*44 λ^- /Cm^R Nm^S, pRK2013 Nm^R::Tn9	(Kessler et al., 1992)
SURE	*endA*1 *glnV*44 *thi*-1 *gyrA*96 *relA*1 *lac recB recJ sbcC umuC*::Tn5 *uvrC* e14- Δ(*mcrCB-hsdSMR-mrr*)171F'[*proAB*$^+$ *lacI*^{q} *lacZ*ΔM15 Tn10]	Stratagene
S17-1	*thi pro hsdR recA* chromosomal RP4 [Tra+ Tc^S Km^S Ap^S]	(Simon et al., 1983)
ST18	*thi pro hsdR recA hemA* chromosomal RP4 [Tra+ Tc^S Km^S Ap^S]	(Thoma and Schobert, 2009)
Plasmids		
pJQ200KS	Suicide vector, p15A ori, $Gent^R$	(Quandt and Hynes, 1993)
pJQ200KSΔP*lac*	pJQ200KS, Δ*lac* promoter	This study
pME3087	Suicide vector, ColE1 ori, Tet^R	(Voisard et al., 1994)

Table S2. List of primers used in this study.

Primer	**Sequence**
BRIP_sylC_KO_P1	TAT GGA TCC GCT CAG GAA ACA GCG ACT TC
BRIP_sylC_KO_P2	ATC AGT CGA CCC GTA CCC TTG ACC AGA CTC
BRIP_sylC_KO_ori	CAG GCT GAA CCA CCA ACT G
gacS_new_P1	TAT GGA TCC AGA ACT GCA ACT GAG CAT CG
gacS_new_P2	ATC AGT CGA CCT GAA GGC CTG GAA CAG TG
gacS_new_ch1	CTC GAT CAG GCG TTT GGA G
hrcC_ins_P1	TAT GGA TCC AGA GCC CGG AAG AGT TTC TC
hrcC_ins_P2	ATC AGT CGA CTG GAC AGT TCT GCC AGC TC
hrcC_ori_ch1	GAG GCT GAG CCG TTA CCT TC
hrpL_new_P1	TAT GGA TCC GAA ACC GCT ACG CCA TAC AG
hrpL_new_P2	CAC TCA GGC GAA GAT CAC AAG ATT CGG GAG CAT
hrpL_new_P3	GAA TCT TGT GAT CTT CGC CTG AGT GAA CAT CTG
hrpL_new_P4	ATC AGT CGA CTC AAG GGA TTG AGG ATC AGC
hrpL_new_ori_A	GTA ATC GCC GAA AGA TCG TC
hrpL_new_ori_B	TTT AAT CAG CGC TTT CAG ACC
hopBA_del_P1	TAT GGA TCC CTC GGC TTT CAT GTC TTG C
hopBA_del_P2	TTA TCC TCG CCA TTT AGC ATA TTA TGT CCT TCG TTG
hopBA_del_P3	ATA TGC TAA ATG GCG AGG ATA ACT AAA GAA AAC C
hopBA_del_P4	ATC AGT CGA CTT GCC AAG GGT GAT GTA GTG
hopBA_oriA	TTA TAA CCC GCG AGA GTT GC
hopBA_oriB	GGT TTG CTG GAC TGG CCT AC
hopA_del_P1	TAT GGA TCC CTG CTA TCC CGG TCA TCG
hopA_del_P2	TAC TCA TTC CAA TAT CGG GTT CAT GAC TCA TCC TC
hopA_del_P3	TGA ACC CGA TAT TGG AAT GAG TAT GAA CTG CTG
hopA_del_P4	ATC AGT CGA CCC AGA ATC TCC AGC TTA CCG
hopA_oriA	AAT CAA GGG CGG CAT AGA C

hopA_oriB	CAT CGA TCA GAA TCG ACG AC
hopAZ-S_del_P1	TAT GGA TCC GTT GTG ACG CAC GAA GTA GC
hopAZ-S_del_P2	CAT CGC GCC ATA AAG GTC ATG GGT ATT TTC CTC
hopAZ-S_del_P3	CCA TGA CCT TTA TGG CGC GAT GGA AGT AAG
hopAZ-S_del_P4	ATC AGT CGA CCT GTA GCT GCC GTA GCA GTC
hopAZ-S_oriA	ACG AAC GCG GTG ATA TTG AC
hopAZ-S_oriB	TCC TCT GTG CTG ACT TGT CG
LscA_del_P1	TAT GGA TCC ATA CGT TGC AGG CAG GTT G
LscA _del_P2	CAG CTC AGC TGC ACA TTG ATG TTA CTC ATG ATG ATG ATG
LscA _del_P3	TGA GTA ACA TCA ATG TGC AGC TGA GCT GAC AGG
LscA _del_P4	ATC AGT CGA CGC CAG TCG TTG AAG GAG AAC
LscA _del_oriA	ATG AGC GAC GTC TTT GTC G
LscA _del_oriB	GCA GCA ACT CGA CCT GTA CC
LscC_del_P1	TAT GGA TCC CCA TGA TGA GAG AAC CTC ACG
LscC _del_P2	TCA GCT CAG TTG CAC GCT GCT ATT GGA CAT AAT TGA TAC
LscC _del_P3	TGT CCA ATA GCA GCG TGC AAC TGA GCT GAC ACG
LscC _del_P4	ATC AGT CGA CGG CAT CAC TGG TGT TCA TTG
LscC _del_oriA	CCG AGG ACG AAT CAT ACC C
LscC _del_oriB	CTT CGC TGA TGT TCC TGT CC
PslD_del_P1	TAT GGA TCC AAC TAC TGC AAT GGC GAA GC
PslD_del_P2	CAA TTA ATT GTT GTT GCG TAA GCT GGC GAT CAG CAT C
PslD_del_P3	GAT CGC CAG CTT ACG CAA CAA CAA TTA ATT GTC AGG
PslD_del_P4	ATG GCT CGA GAC CGG TCG TCT TCA CTT CC
PslD_del_oriA	GAA AAA GTC GCC GTG GTG
PslD_del_oriB	TAC TTA CCG CGC TCA ACT GG
AlgA_del_P1	TAT GGA TCC CCT TCG TGC GTG TCT ACA AC
AlgA_del_P2	CTG TCA GCG AGC TGG AAT CAT TTT GTT TCT CCA AAA G
AlgA_del_P3	CAA AAT GAT TCC AGC TCG CTG ACA GAC GCT TTT AG
AlgA_del_P4	ATC AGT CGA CGG CTG TAA ATG GCA GAT TCC
AlgA_del_oriA	ACG CCA AAC AAT CCG AGT C
AlgA_del_oriB	AAA CCG AAG AGG GAT TGA CC
PssB64_0893_KO_P1	TAT GGA TCC TGC TGC AAT TGC TCG ATA AC
PssB64_0893_KO_P2	ATG GCT CGA GCT TCG ATG ACC GAC TTG AGG
PssB64_0893_ch1	GAT CAG CGT GTG GGC TTC
PssB64_3440_KO_P1	TAT GGA TCC AGA CCT GAT GGA ATG CTT GC
PssB64_3440_KO_P2	ATG GCT CGA GAA TTA CGC CCT TCA AAC GAC
PssB64_3440_ch1	CTG CAA TTG TCC GTA ACT CG
B64_3440_KO_EcoRI	ATC CGG AAT TCA ATT ACG CCC TTC AAA CGAC
pJQ200KS_B1_R	TCT AGA ACT AGT GGA TCC
pr_3087 H3	GAG AAA TCA CCA TGA GTG

Table S3. List of the mutated and other genes mentioned in this work with the corresponding locus tags.

Strain	Gene	Locus tag	Mutant type
Psy SM	*hrcC*	PssSM_1194	Knock-out*
	hrpL	PssSM_1210	Deletion#
	gacS	PssSM_3753	Knock-out
	hopBA1	PssSM_2782	Deletion
	hopAZ1	PssSM_1866	Deletion
	hopA2	PssSM_1213	Deletion
	lscA	PssSM_2171	Deletion
	lscC	PssSM_0731	Deletion
	algA	PssSM_1040	Deletion
	pslD	PssSM_3336	Deletion
	clpV-1	PssSM_2483	Knock-out
	clpV-2	PssSM_5057	Knock-out
	avrE‡	PssSM_1182	-
	hopM1‡	PssSM_1180	-
	hopI1‡	PssSM_4439	-
	hopAA1‡	PssSM_1177	-
	tssA-1‡	PssSM_2478	-
	tssA-2‡	PssSM_5052	-
Psy B64	*sylC*	PssB64_04153	Knock-out
	hrcC	PssB64_04549	Knock-out
	gacS	PssB64_02187	Knock-out
	clpV-1	PssB64_00893	Knock-out
	clpV-2	PssB64_03440	Knock-out
BRIP34876	*sylC*	A979_11059	Knock-out

* Inactivation by plasmid insertion

In-frame deletion of the sequence

‡ No mutants of these genes were generated

7. Genomics-Based Exploration of Virulence Determinants and Host-Specific Adaptations of *Pseudomonas syringae* Strains Isolated from Grasses

Alexey Dudnik[1] and Robert Dudler[1,2]

Pathogens

3(1):121-148

doi: 10.3390/pathogens3010121

Published: 28 January 2014

[1] Institute of Plant Biology, University of Zurich, Zollikerstrasse 107, CH-8008 Zurich, Switzerland

[2] Corresponding author

7.1. Abstract

The *Pseudomonas syringae* species complex has recently been named the number one plant pathogen, due to its economic and environmental impacts, as well as for its role in scientific research. The bacterium has been repeatedly reported to cause outbreaks on bean, cucumber, stone fruit, kiwi and olive tree, as well as on other crop and non-crop plants. It also serves as a model organism for research on the Type III secretion system (T3SS) and plant-pathogen interactions. While most of the current work on this pathogen is either carried out on one of three model strains found on dicot plants with completely sequenced genomes or on isolates obtained from recent outbreaks, not much is known about strains isolated from grasses (Poaceae). Here, we use comparative genomics in order to identify putative virulence-associated genes and other Poaceae-specific adaptations in several newly available genome sequences of strains isolated from grass species. All strains possess only a small number of known Type III effectors, therefore pointing to the importance of non-Type III secreted virulence factors. The implications of this finding are discussed.

7.2. Introduction

Strains of the *Pseudomonas syringae* group are causal agents of a variety of plant diseases worldwide. Strains of this group have been reported to infect nearly 200 different plant species (Bradbury, 1986), including both grassy and woody hosts. A number of those are agriculturally important plants, and as a consequence, *P. syringae* is one of the best-studied plant pathogens. While some diseases have been known to recur in the form of outbreaks for a long time (Arnold et al., 2011; Gironde and Manceau, 2012; Martín-Sanz et al., 2013; Ramos et al., 2012), others have only recently emerged (Green et al., 2010; McCann et al., 2013). The taxonomy of *P. syringae* is still under discussion. According to the commonly used version, the species is sub-divided into over 50 groups called “pathovars”. This division takes into account host-specificity, disease type, as well as the biochemical characteristics of a strain (Qi et al., 2011; Young, 2010). In contrast, a number of phylogenetic studies have positioned *P. syringae* as a species complex, and depending on the approach used, it has been split either into five phylogenetic groups by multi-locus sequencing (Hwang et al., 2005; Sarkar and Guttman, 2004), or into nine so-called genomospecies based on

DNA hybridization profiles (Gardan et al., 1999). In order to remain coherent with the majority of publications, in this work, we are going to regard *P. syringae* as a single species.

The observed large genetic diversity among different pathovars is a direct consequence of the wide host range. Moreover, heterogeneity is also found among strains belonging to the same pathovar (McCann et al., 2013; O'Brien et al., 2012; Qi et al., 2011). The highest degree of variation is seen within the complement of virulence factors, which is the key element determining host range and the overall degree of virulence (Lindeberg et al., 2009). In addition to that, some strains possess host-specific metabolic pathways, such as enzymes for lignin degradation in pathogens of woody hosts (Green et al., 2010; McCann et al., 2013; Rodríguez-Palenzuela et al., 2010). For successful survival and reproduction, both epiphytic and endophytic strains of *P. syringae* deploy different sets of Type III and Type VI secreted effector proteins, exopolymeric substances, phytohormones, phytotoxins and other types of secreted molecules (Bender et al., 1999; Denny, 1995; Haapalainen et al., 2012; Lindeberg et al., 2009, 2012; Misas-Villamil et al., 2013; Rodríguez-Moreno et al., 2008; Schellenberg et al., 2010; Yu et al., 1999). Among those, the major pathogenicity factor is the Type III secretion system (T3SS). Other notable virulence factors are the necrosis-inducing phytotoxins, syringomycin and syringopeptin, which are presumed to create pores in plant cells by imbedding into the plasma membrane, and the anti-metabolite toxins, tabtoxin, mangotoxin and phaseolotoxin, which inhibit glutamine synthetase, ornithine acetyltransferase and ornithine carbamoyl transferase, respectively (Arrebola et al., 2011; Bender et al., 1999). Another group of phytotoxins includes coronatine and syringolin, both of which are involved in the inhibition of the salicylic acid-dependent immune response by mimicking jasmonic acid or by irreversibly inhibiting the proteasome, respectively (Bender et al., 1999; Schellenberg et al., 2010).

The T3SS is a protein delivery machinery, which uses a structure, called injectosome, for the delivery of effector proteins directly into host cells by puncturing the cell membrane (Büttner, 2012; Galán and Wolf-Watz, 2006). This machinery is essential for the pathogenesis of *P. syringae*, and knocking it out renders the bacterium avirulent (Lindeberg et al., 2009, 2012). The majority of Type III effectors (T3Es) are assumed to be involved in the suppression of plant defense, including pathogen-associated molecular pattern (PAMP)-triggered immunity and the hypersensitive response (effector-triggered immunity) (Cunnac et al., 2009; Jones and Dangl, 2006). In addition, some effectors were shown to have a cytotoxic effect (Munkvold et al., 2008; Salomon et al., 2012). Currently, there are 58 verified effector families recognized (Baltrus et al., 2011). With a few

exceptions, the exact mechanism of action of T3Es remains unknown. Among the well-characterized T3Es is AvrPtoB, which targets the flagellin recognition receptor, FLS2, and marks it for degradation (Göhre et al., 2008), HopU1, which ADP-ribosylates several RNA-binding proteins, thus preventing the association with their target mRNAs (Nicaise et al., 2013), HopN1, which targets photosystem II in order to inhibit reactive oxygen species production (Rodríguez-Herva et al., 2012), AvrRps4, which targets a regulator of plant basal defenses (Bhattacharjee et al., 2011), and HopZ1a, which interferes with plant microtubule network formation and jasmonic acid signaling (Jiang et al., 2013; Lee et al., 2012a). Most *P. syringae* strains produce around two to three dozen T3Es (Baltrus et al., 2011; McCann et al., 2013; Qi et al., 2011). However, not all of them are essential for full virulence, due to functional redundancy (Cunnac et al., 2011). As a consequence, strains pathogenic to the same host often have divergent sets of effectors (Almeida et al., 2009; McCann et al., 2013; O'Brien et al., 2012; Qi et al., 2011). It is notable that effector repertoires are under heavy evolutionary pressure (Jones and Dangl, 2006) and, thus, are being continuously remodeled. The remodeling might also result in a change of host specificity. The field of effector biology remains a hot topic, and currently, a lot of research is aimed at improving our knowledge about the molecular biology of interactions between plants and their pathogens.

The extensive research has also led to the accumulation of a large amount of available genome sequence data. Currently, GenBank contains records of the three completely sequenced *P. syringae* model strains DC3000 (pathovar *tomato*, *Pto*; pathogenic to tomato, *Arabidopsis thaliana* and *Nicotiana benthamiana* (Buell et al., 2003; Wei et al., 2007)), B728a (pathovar *syringae*, *Psy*; the causal agent of brown spot disease of bean (Feil et al., 2005)) and 1448A (pathovar *phaseolicola*, *Pph*; causes halo blight on bean (Joardar et al., 2005)). The strains represent phylogenetic Clades one, two and three, respectively (Baltrus et al., 2011; Hwang et al., 2005). In addition, a number of incomplete genome sequences of various qualities are available for a variety of strains (Baltrus et al., 2011; Green et al., 2010; McCann et al., 2013; Qi et al., 2011; Rodríguez-Palenzuela et al., 2010). The majority of the sequenced strains were originally isolated from dicot plants; thus, the topic of diversity and adaptations of strains pathogenic to monocot plants remains largely unexplored. Currently, nine genome sequences of different qualities of strains isolated from true grasses (graminoids, family Poaceae) are available. These encompass four wheat (*Triticum aestivum*) isolates (*P. syringae* pv. *syringae* (*Psy*) strains SM and B64, *P. syringae* pv. *atrofaciens* DSM50255 (*Paf*) and *P. syringae* BRIP39023 (Dudnik and Dudler, 2013a, 2013b; Gardiner et al., 2013)), three strains pathogenic to barley (*Hordeum vulgare*) (*P. syringae* BRIP34876 and

BRIP34881 (Gardiner et al., 2013) and pathovar *japonica* strain M301072 (*Pja*) (Baltrus et al., 2011)), one strain collected from rice (*Oryza sativa*) (pathovar *oryzae* 1_6 (*Por*) (Reinhardt et al., 2009)) and an isolate pathogenic to proso millet (*Panicum miliaceum*) (pathovar *panici* strain LMG2367 (*Ppa*) (Liu et al., 2012)). Here, we analyze these nine genomes using comparative genomics tools with the aim of identifying possible adaptations of *P. syringae* strains to life in graminoid host species. In addition, we also compared these strains with a group of unspecialized pathogenic strains belonging to *P. cannabina* pv. *alisalensis* (*Pcal*), which have been shown to colonize several dicot hosts, as well as oat (*Avena sativa*) and great brome (*Bromus diandrus*) plants, both of which belong to the Poaceae (Sarris et al., 2013).

7.3. Results and Discussion

7.3.1. Phylogenetic assessment of the strains

Since most of the analyzed genomes have been sequenced within the past year, there is no published record of their phylogenetic characterization. Thus, the first step undertaken was to identify their relatedness by constructing a maximum likelihood tree using MLST (Multi-locus sequence typing) loci previously deployed in other studies(Baltrus et al., 2011; Sarkar and Guttman, 2004). Several other strains belonging to the three major phylogenetic clades (Hwang et al., 2005) have been included as a reference. *P. fluorescens* BRIP34879, a strain which was also isolated from barley (Gardiner et al., 2013), was used as an outgroup. *Pcal* ES4326 (previously known as *P. syringae* pv. *maculicola* ES4326) was included to represent the phylogenetic relatedness of the *P. cannabina* pv. *alisalensis* group. The resulting diagram is presented in Figure 1. Even though whole-genome-based phylogenies were recently demonstrated to be more accurate than MLST (Baltrus et al., 2013), we still decided to use the MLST approach, due to the poor assembly quality of some of the analyzed genomes, which resulted in a large number of partial and split genes. The presence of such sequences would introduce a bias to the analysis using the whole genome/proteome sequences.

While most of the *P. syringae* strains that are the focus of this article cluster together within Clade II, *Por* 1_6 clearly forms a separate branch. This is in line with previously published data, where this strain was placed into Clade IV (also classified as *Pseudomonas coronafaciens* by Gardan and colleagues (Gardan et al., 1999)). Interestingly, Clade IV contains several other strains isolated from cereal crops, including oat (Hwang et al., 2005). This suggests that a host-shift towards Poaceae has occurred at least twice in the evolution of *P. syringae*. Within Clade II, strain BRIP39023, a non-

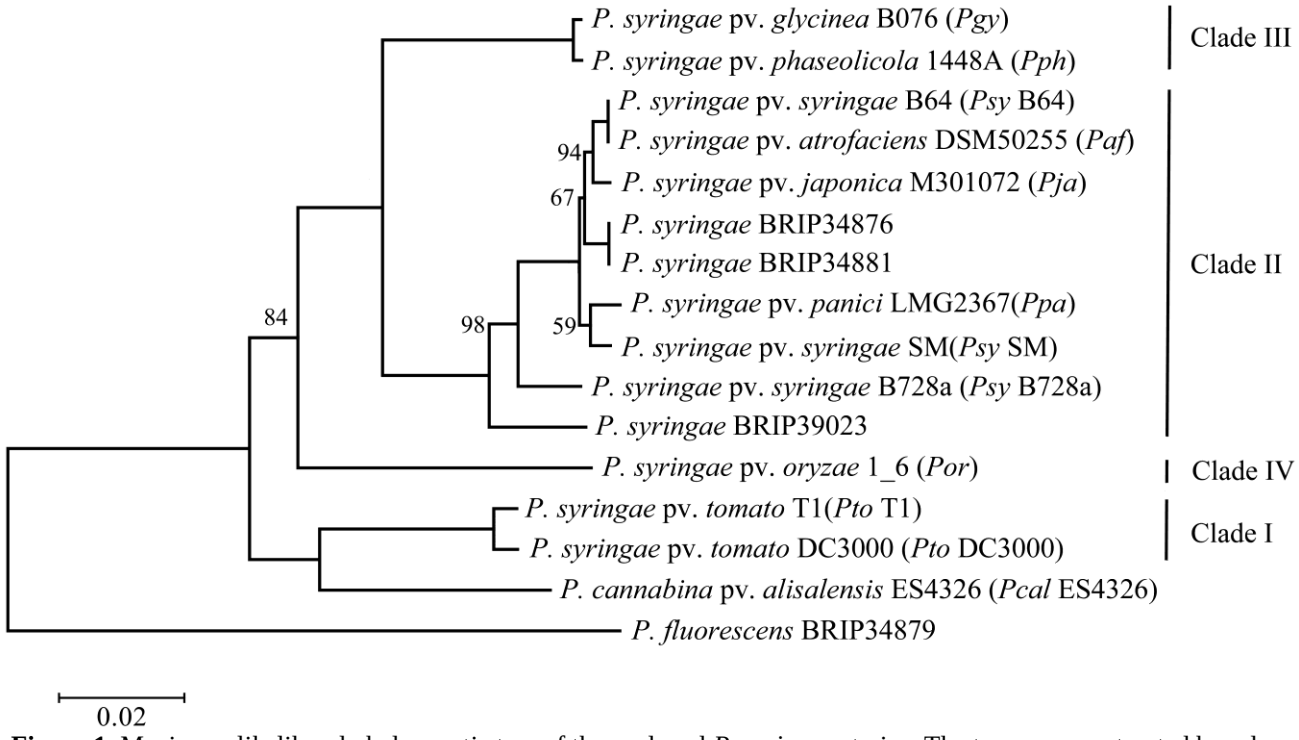


Figure 1. Maximum likelihood phylogenetic tree of the analyzed *P. syringae* strains. The tree was constructed based on the nucleotide sequences of several housekeeping genes. The scale bar indicates the number of amino acid changes per site. Bootstrapping-based probabilities are displayed only at nodes where the values are less than 100.

pathogenic wheat isolate, forms a separate branch and, therefore, is also less related to the rest of the isolates. The remaining strains form a cluster and, therefore, must be of the same ancestral lineage. It is notable that pathovars *syringae*, *atrofaciens*, *panici* and *japonica* do not form distinct branches. This shows that discrimination into pathovars is not always descriptive, since particular strains often are pathogenic to several hosts, as shown, for example, for strains BRIP34876, BRIP34881 and *Ppa* (Gardiner et al., 2013; Liu et al., 2012). This also suggests that a pathogen can relatively easily change its specificity from one host to another among related plant species. Finally, it is not clear to which pathovar BRIP34876 and BRIP34881 should be assigned. It should be noted that there are several other *P. syringae* pathovars belonging to Clade II (*lapsa*) and Clade IV isolated from Poaceae (Elasri et al., 2001; Gardan et al., 1999), which were not included in this study, due to the lack of available genomic sequence data. *Pcal* ES4326 forms a distinct branch from *Pto* strains and, together with several related strains, was classified as Clade V by Hwang and colleagues (Hwang et al., 2005). The resulting tree is in agreement with previously published data (Baltrus et al., 2011; Hwang et al., 2005).

7.3.2. Genome comparison and identification of Poaceae-specific genes

During the process of adaptation to a new host, there is a strong evolutionary pressure on a pathogen. As a consequence, a pathogen loses some genes that reduce the virulence and overall fitness in the new host. In addition, novel genes are acquired by horizontal gene transfer (HGT), in particular from species already living on or inside the new host. Examples include genes for the degradation of lignin, pectin and aromatic compounds found in *P. syringae* strains pathogenic to woody hosts (Marcelletti et al., 2011; Rodríguez-Palenzuela et al., 2010). Therefore, the strains isolated from grasses could also exhibit certain adaptations not found in other strains. In order to identify candidate adaptation genes, the genomes of the nine sequenced Poaceae isolates were first compared among themselves and, then, also, with a group of twelve other *P. syringae* strains isolated from a variety of hosts (See Table S1 for a complete list of strains). The comparison was performed using an 80% identity cut off in order to avoid the detection of paralogs and other genes with only partial homology. An outline of the results is shown in Table 1, and a list of the identified unique genes, which are shared by at least four strains, can be found in Table S2.

Table 1. Genome comparison of the nine analyzed Poaceae isolates: unique genes.

Ortholog clusters	A (within grass isolates)	B (not found outside grass isolates)
Unique to *Por*	2,333	1,566
Unique to BRIP34876	11	8
Unique to BRIP34881	13	7
Unique to BRIP39023	349	121
Unique to *Pja*	3,657	2,563
Unique to *Psy* SM	178	75
Unique to *Psy* B64	37	9
Unique to *Ppa*	484	322
Unique to *Paf*	216	188

Column **A** represents the numbers of ortholog clusters obtained by comparing the nine genomes of Poaceae isolates among themselves. Column **B** contains an overview on the distribution of ortholog clusters found exclusively in these nine isolates, but not outside. An ortholog cluster is a group of genes from at least one strain in which all members have an identity percentage equal to or above the set cut off. Thus, a single cluster might contain more than one gene per strain, due to the presence of recent gene duplications, which are still more than 80% identical to one another. Such genes are regarded as the same entity by the software, and as a consequence, the actual numbers of shared homologs are slightly different for each strain.

When the Poaceae isolates were compared to one another, the total number of identified ortholog clusters was 13,319. The calculated core genome, *i.e.*, the set of genes shared by all strains, has a size of 3,578 ortholog clusters (Figure 2A), which is comparable to the previously published data (Baltrus et al., 2011). Outside the core genome, 2,471 clusters are shared by at least two strains. An overview of shared genes is depicted in Figure 2A. The remaining ortholog clusters have only one member, and respective genes are therefore unique to one of the strains. The number of unique genes is highly variable, ranging from as little as eleven for BRIP34876 to almost 3,657 for *Pja* (Table 1, Column A). The large values observed for *Pja* and *Por*, however, originate from the low quality of the respective assemblies (see Table S1). Because of this, there are many proteins that are found on several contigs. For example, the *Pja* genome contains seven entries for the chromosomal replication initiation protein, DnaA, which is normally present as a single copy next to the origin of replication. It is highly likely that many partial sequences did not pass our strict identity criterion and therefore ended up as individual ortholog clusters. Moreover, the majority of the genomes did not undergo manual curation, and thus, their annotation data contains a number of very small hypothetical, as well as truncated proteins.

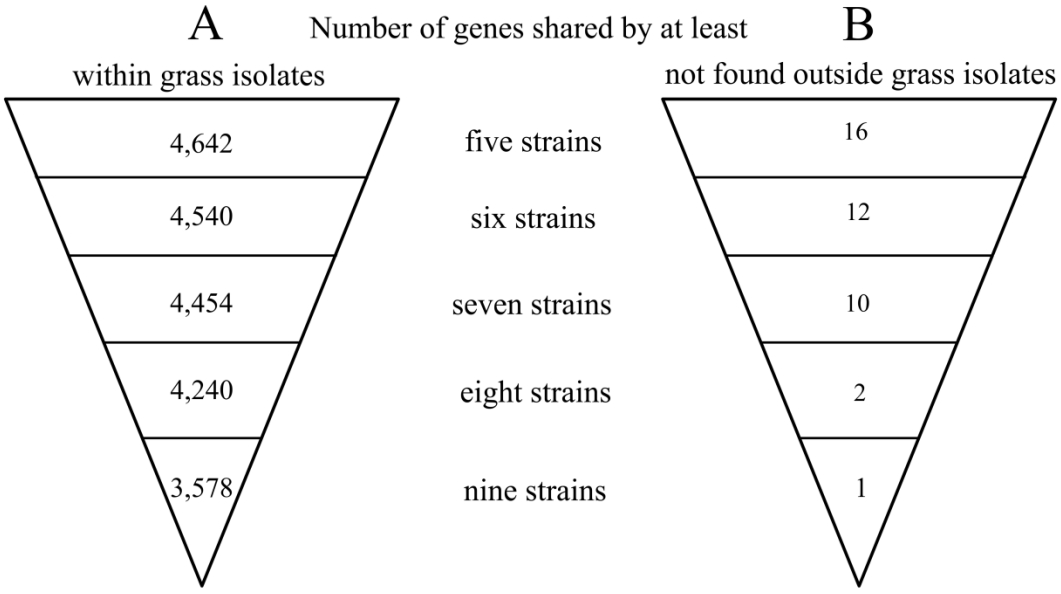


Figure 2. Shared genes in the genomes of the nine analyzed Poaceae isolates. Column **A** represents the numbers of ortholog clusters shared by five or more of the nine Poaceae isolates when compared among themselves. Column **B** contains an overview of the distribution of ortholog clusters found exclusively in these isolates, *i.e.*, ortholog clusters that were not detected in any other analyzed *P. syringae* strain.

Surprisingly, however, there is only a single gene that is found in all nine Poaceae isolates, but is absent in any of the other *P. syringae* genomes included in the comparison (Figure 2B and Table S2;

see Section 2.4.2). Furthermore, the additional gene shared by eight strains (all except BRIP39023) encodes a methyltransferase, which has no homologs in any other sequenced genome. Moreover, there are only 36 additional genes that are found in at least four of the Poaceae-colonizing strains (Figure 2B and Table S2). Out of these 36, only four are present in the *Por* genome, which is most likely due to the different phylogenetic lineage of this strain. Interestingly, most of these genes are found within four genomic regions that are flanked by sequences with different G + C profiles and transposable elements (as marked in Table S2). Regions 1 and 2 are present in all of the Clade II strains, except BRIP39023, which makes it very tempting to speculate that the asymptomatic phenotype of this strain could be, at least partially, due to the absence of some of these genes. This is not completely unjustifiable, as these genomic regions encode several proteins that could potentially be involved in the protection from oxidative stress and anti-microbial compounds. These include a catalase-like protein, cytochrome b561 and a small multidrug resistance protein. Additionally, there are also three transcriptional regulators that might be needed for the fine-tuning of gene regulation. The other two genomic regions are present in only four strains, but aside from hypothetical proteins, they encode a putative formaldehyde dehydrogenase, a putative general stress protein, a sensor histidine kinase and an MFS (Major Facilitator Superfamily) transporter. However, this conclusion is purely speculative, and the characterization of deletion mutants of these regions is required to elucidate their function.

7.3.3. The type III secretion system and effector repertoire

Based on genome sequences, the analyzed strains possess the canonical *hrp*/*hrc*-type type III secretion system (Büttner, 2012; Deng et al., 1998), which is found in most of the sequenced *P. syringae* isolates. The effector repertoires for each strain are presented in Figure 3. Based on the number of T3Es, the strains form two distinct groups: the first one consists of *Por*, which has 27 full-length T3Es, while the remaining eight strains only contain around a dozen effectors. The distribution also correlates with the phylogenetic relatedness of the strains (see Figure 1), further supporting the assumption that the host shift has occurred independently for the two groups. Moreover, *Por* has only five effectors shared with *Ppa*, which was also reported to infect rice (Liu et al., 2012), therefore supporting the idea that the same host can be colonized by strains with very little overlap in effector composition. Interestingly, both strains lack AvrE1. Unfortunately, however, there are no reliable data on T3Es from other rice isolates to further validate these findings.

Apart from *Por*, the remaining eight *P. syringae* strains contain a clear core effector set consisting of AvrE1, HopI1, HopAA1, HopM1 and HopBA1, of which the first four are found in intact form in most sequenced *P. syringae* strains (Baltrus et al., 2011; McCann et al., 2013). The phylogenetic relationship of the core T3Es is shown in Figure S1. The presented individual phylogenetic trees in general reflect the one obtained using the MLST approach (Figure 1). Interestingly, HopBA1 is identical in nearly all strains, suggesting a strong evolutionary pressure to maintain the sequence. Moreover, the remaining effectors appear to have different substitution rates, as seen from the corresponding scale bars. The other most common effectors are HopZ3, HopAZ1 and the HopAG1-HopAH1-HopAI1 cluster. The effector repertoire also correlates with phylogenetic data, as, for example, *Ppa* differs from its closest relative, *Psy* SM, only by the HopAG1-HopAH1-HopAI1 locus. Interestingly, the avirulent strain, BRIP39023, has the largest number of effectors, implying that the number of effectors does not correspond to disease severity and overall virulence. However, it is also possible that one of the effectors acts as an avirulence (Avr) protein, triggering a local hypersensitive response and, therefore, suppressing endophytic growth.

All five currently sequenced *Pcal* strains contain a relatively large number of T3Es, which includes 28 to 32 full-length, as well as five to eleven truncated or disrupted effectors (Sarris et al., 2013). Out of those, 16 full-length T3Es are shared with *Por*: AvrPto1, HopAA1, HopAB3, HopAD1, HopAO1, HopAQ1, HopAS1, HopBF1, HopD1, HopG1, HopM1, HopQ1, HopR1, HopV1, HopBI1 and HopX1. So far, all *P. syringae* strains isolated from oat that were used in scientific studies belong to Clade IV (Hwang et al., 2005; Sarkar et al., 2006) and, thus, are likely to have comparable effector repertoires. Assuming that this is true and applies also to the *P. syringae* pv. *coronafaciens* (Clade IV) strains, one might expect that they might be able to infect both monocots and dicots, too. However, these strains were reported to cause disease symptoms on brome, rye and oat, but not on crucifers or tomato (Cintas et al., 2002). Interestingly, neither *Pcal* nor pathovar *coronafaciens* strains were able to cause disease in wheat (Ishiyama et al., 2013). In contrast to *Por*, *Pcal* strains have very little overlap with the Clade II *P. syringae* strains with regard to their T3E repertoire. In addition to the core effectors, AvrE1, HopAA1, HopM1 and HopI1, only HopAL1, HopAZ1 and HopBF1 are found in some *Pcal* strains.

Having relatively few T3Es is not uncommon for *P. syringae* and seems to be a feature of Clade II strains (Baltrus et al., 2011). However, the strains isolated from Poaceae species are rather extreme examples, and at present, it remains unclear why this is the case. Some T3Es form so-called

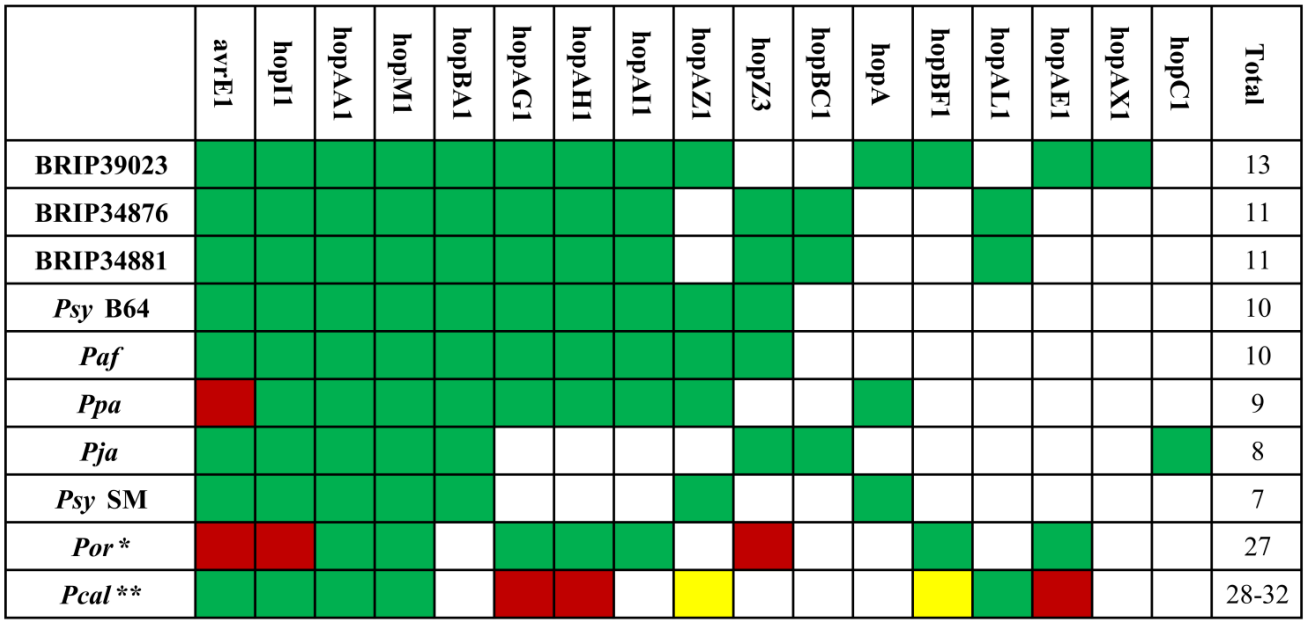


Figure 3. Type III effector repertoires. Green squares indicate the presence and blank squares indicate the absence of the respective T3Es, and red squares indicate a truncated or disrupted effector. Yellow indicates that an effector is not present in all strains assigned to this species. The total number of effectors is indicated at the end of each row. * The genome of *Por* contains an additional 20 full-length effector-coding genes not found in any of the eight other analyzed *P. syringae* strains (*hopAQ1*, *avrB1*, *avrB3*, *avrPto1*, *avrPto4*, *hopAD1*, *hopAS1*, *hopX1*, *hopAO1*, *hopD1*, *hopG1*, *hopR1*, *hopV1*, *hopY1*, *hopAB3*, *hopH2*, *hopQ1-1*, *hopQ1-2*, *hopBH1* and *hopBI1*). ** Likewise, for the *Pcal* strains, only effectors that are also present in Clade II strains are shown. For a detailed T3E composition of the individual *Pcal* strains, see (Sarris et al., 2013). It should also be note that some of the genomes of these strains encode a number of the so-called discontinued T3Es, which were never experimentally shown to be translocated in a T3SS-dependent manner: HopAN1, HopAC1, HopJ1, HopL1 and HopAH2.

Redundant Effector Groups (REGs) (Cunnac et al., 2011), where members of a REG target different steps in a particular defense signaling pathway. Thus, effectors within one REG are often interchangeable, and the possession of one member of each of the required REGs should be sufficient for full virulence. In support of this notion, the minimal effector repertoire for *Pto* DC3000 on *A. thaliana* has been identified to consist of eight out of 31 T3Es (Cunnac et al., 2011), a value which is similar to what is seen in the Poaceae-colonizing strains. Therefore, it is possible that in these hosts, there is a strong selection pressure to maintain just one effector for each REG. On the other hand, there is always the possibility that these genomes contain a number of unrecognized novel T3Es. However, based on results by Baltrus and colleagues (Baltrus et al., 2011), there should be very few undiscovered effector families, and it is therefore unlikely that our search missed a considerable number of unknown T3Es.

In addition to the *hrp*/*hrc*-type T3SS, several *P. syringae* strains were described to possess a second rhizobial-like T3SS, belonging to the *Rhc* family subgroup II (Gazi et al., 2012). This T3SS-2 was originally identified in *Pph*, *Pta* and *Por*. In *Por*, the gene cluster encompasses 31 genes (POR16_18228–POR16_18383) and appears to contain all essential structural components (Gazi et al., 2012). Among the other Poaceae isolates, we have identified another T3SS-2 gene cluster in BRIP39023 (A988_16153–A988_16283, accession no. KB316298). Structure-wise, the gene cluster rather resembles its homolog in *Pta* than that of *Por* or *Pph*. It is not, however, known whether this T3SS is active and plays any role in interactions with the host. Moreover, the genomes of both BRIP39023 and *Por* do not contain any known rhizobial effectors, based on sequences from *Rhizobium*, *Sinorhizobium* and *Bradyrhizobium* deposited on the T3DB (Type-III-Secretion-System Related Database) website (Wang et al., 2012).

7.3.4. Other virulence factors

7.3.4.1. Phytotoxins and other small secreted molecules

While the major pathogenicity determinant is the T3SS, many *P. syringae* strains also rely on other secreted molecules for entry and colonization of their hosts. The most important and best-studied group of such compounds are the phytotoxins. There seems to be no general requirement for phytotoxin production, and while some strains are capable of producing several different compounds, others lack any known phytotoxins (Baltrus et al., 2011; Bender et al., 1999). Moreover, even strains belonging to the same pathovar have differences in phytotoxin gene composition (Cai et al., 2011; Murillo et al., 2011; Qi et al., 2011). Nevertheless, it is a well-established fact that their presence enhances disease progression and symptom development (Arrebola et al., 2009; Bender et al., 1999; Groll et al., 2008). In terms of overall distribution, syringolin, syringomycin, syringopeptin and mangotoxin are found almost exclusively among Clade II strains, while all other phytotoxins are found outside of this phylogenetic group (Baltrus et al., 2011; Carrión et al., 2013). So far, the only described exception is *P. syringae* pv. *syringae* CFBP3388, a Clade II strain that was identified to produce both phaseolotoxin and mangotoxin (Carrión et al., 2013; Murillo et al., 2011).

The composition of phytotoxin biosynthetic gene clusters among the Poaceae isolates (Figure 4) fits into the above-mentioned phylogenetic clade-dependent distribution scheme. The genome of *Por*, which belongs to Clade IV, contains gene clusters for the biosynthesis of coronatine and tabtoxin,

whereas all other strains have genes for the production of mangotoxin, syringopeptin and syringomycin. However, *Psy* SM seems unable to produce syringomycin, as its *syrE* gene is truncated. It should be noted that the biosynthetic gene clusters for syringopeptin and syringomycin encode very large non-ribosomal peptide synthetases (NRPS) containing repetitive sequences. Because these clusters are often found on several contigs, their complete intactness cannot be assessed with certainty. Complete and intact syringolin synthetase gene clusters, and, thus, the ability for syringolin production, are present in all Clade II strains analyzed, except in *Psy* SM, *Ppa* and BRIP39023. The first two do not have a syringolin synthetase gene cluster, whereas the latter strain contains a frameshift mutation in the *sylA* gene. Genomes of the *Pcal* strains only contain coronatine biosynthesis genes. None of the analyzed strains is capable of producing phaseolotoxin. In conclusion, there appears to be no specific requirement in terms of toxin composition for being able to colonize Poaceae species.

Some *P. syringae* strains are known to produce phytohormones, such as the auxin indole 3-acetic acid (IAA), cytokines and ethylene (Akiyoshi et al., 1987; Costacurta and Vanderleyden, 2008; Weingart et al., 1999). While cytokine biosynthesis appears to be specific for *P. syringae* pv. *savastanoi* and ethylene production was detected only in pathovars *pisi*, *cannabina*, *glycinea*, *phaseolicola* and *sesame*, IAA biosynthetic genes appear to be widespread among all major phylogroups (Baltrus et al., 2011; Weingart et al., 1999). Interestingly, none of the Poaceae isolates have homologs of the characterized tryptophan 2-monooxygenase (*iaaM*) and indoleacetamide hydrolase (*iaaH*) genes from *P. syringae* pv. *savastanoi* (*Psv*) NCPPB3335 (accession M11035). However, there appears to be an alternative pathway for IAA biosynthesis, as there are a number of strains where IAA production was demonstrated experimentally, but which lack homologs of the $iaaM^{Psv}$/$iaaH^{Psv}$ genes(Glickmann et al., 1998). The majority of *P. syringae* genomes, including the nine Poaceae isolates, encode another gene annotated as tryptophan 2-monooxygenase at a different genomic location (see accession AAY39694 as a reference). There is also an R-amidase-like protein encoded 15 bp downstream of this gene, and most likely, these two genes constitute an operon. Considering the fact that IaaH also belongs to the amidase family, it is possible that this putative operon is responsible for IAA production in strains lacking the $iaaM^{Psv}$/$iaaH^{Psv}$ homologs. A similar situation is observed for the *Pcal* strains. Apart from that, N-ε-(indole-3-acetyl)-L-lysine synthetase (*iaaL*), which modifies IAA, is present in both the *Por* and *Pcal* strains (Roberto et al., 1990). The ethylene biosynthesis gene, *efe*, is present only in the genomes of the *Pcal* strains, but not in the Poaceae isolates.

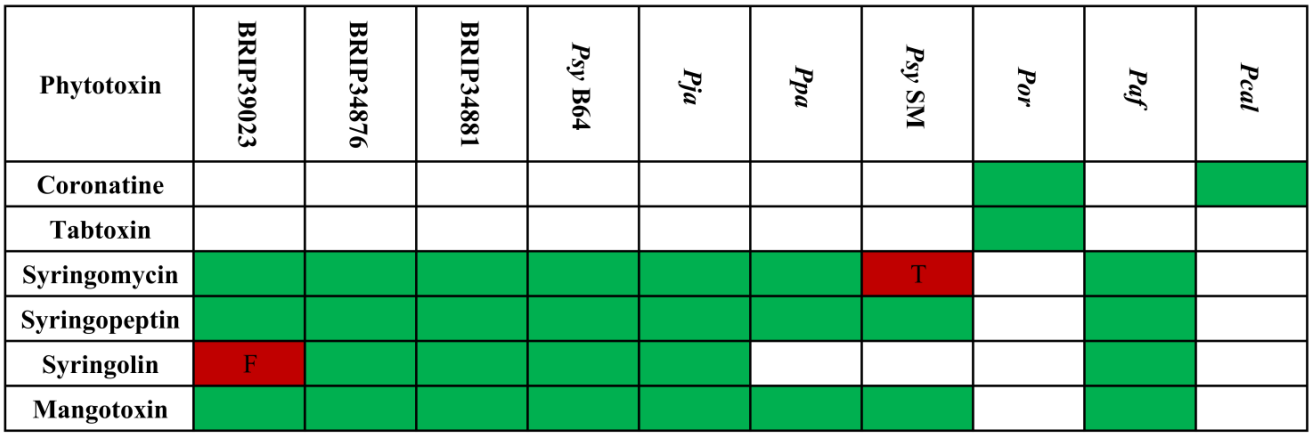


Figure 4. Distribution of phytotoxins among the analyzed strains. Green squares indicate the presence, while blank squares indicate the absence of the respective biosynthetic genes. "F" indicates a frameshift in at least one gene, "T" indicates that at least one gene is truncated.

7.3.4.2. Quorum sensing

Bacterial cell-to-cell communication is an important part of bacterial physiology that influences many processes, such as growth, differentiation and virulence (Miller and Bassler, 2001). One type of such intercellular communication is dependent on cell density and referred to as quorum sensing (QS). The first described QS system is regulating luminescence in *Vibrio fischeri* (Nealson, 1977), and ever since its discovery, QS research continues to be a hot topic, due to its large spectrum of influence. A typical Gram-negative type QS system consists of two proteins encoded by adjacent genes: a LuxI-family synthetase, which produces a signaling compound (in the vast majority of cases, these are acyl-homoserine lactones, AHL) and a LuxR-family transcriptional regulator. Upon perception of the respective AHL, the LuxR protein activates transcription of *luxI*, as well as of other genes that contain a *lux*-type box in their promoters (Miller and Bassler, 2001). QS has been show shown to be important for a number of plant pathogens, including *Agrobacterium tumefaciens*, *Pantoea stewartii*, *Xanthomonas campestris* and *Erwinia carotovora*, where it regulates processes, such as Ti plasmid conjugation, secondary metabolite production, enzyme and exopolysaccharide (EPS) secretion, *etc.* (Von Bodman et al., 2003). Many *P. syringae* strains also possess a QS system encoded by the *ahlR-ahlI* locus (Dumenyo et al., 1998), which was demonstrated to be involved in the regulation of epiphytic fitness (Von Bodman et al., 2003).

Interestingly, none of the analyzed Poaceae-colonizing strains, as well as none of the currently sequenced *Pcal* strains contain *ahlR-ahlI* homologs or any other known quorum sensing system found in Gram-negative bacteria. The trend seems to be general for strains isolated from Poaceae

and also extends to pathovars *atrofacies* and *lapsa* (Elasri et al., 2001). Moreover, strain *Ppa* CFBP 2345 used in that study and *Ppa* LMG2367 are identical (http://goo.gl/EQLjo4), thus validating our genomics-based findings. Because of the demonstrated importance of the QS for a number of pathogens, its loss might be expected to have negative consequences on leaf colonization. However, since the lack of the canonical QS system is relatively abundant among *P. syringae* group strains (Elasri et al., 2001), it is likely that the AHL-mediated gene regulation has no significant effect on the T3SS in these strains. As to why the lack of QS appears to be a general pattern for Poaceae isolates, this could be connected to the fact that some groups of AHLs are actually recognized by plant cells and induce changes in gene expression, including defense genes (Venturi and Fuqua, 2013). Moreover, exposure to AHL was shown to be an inducer of systemic acquired resistance in barley (Schikora et al., 2011). The AHL type used in that study was different from the one produced by *P. syringae*. However, it is likely that the recognition spectrum is broader, thus making AHL producers less virulent and giving a selective advantage to those which have lost the respective genes.

Bacterial genomes often encode orphan LuxR-type transcriptional regulators, which are not associated with a cognate LuxI partner. Some were shown to bind signaling molecules originating not only from cells of the same species, but also coming from other bacteria and even from eukaryotic organisms (Brencic and Winans, 2005; Subramoni and Venturi, 2009; Venturi and Fuqua, 2013). There also exists a distinct clade of the orphan LuxR proteins identified in several groups of plant-associated bacteria, which were reported to have an affinity towards plant-borne low-molecular weight compounds (Subramoni et al., 2011). For example, the LuxR family member OryR from *X. oryzae* pv. *oryzae* was identified to be a global regulator and a virulence factor that responds to the presence of homogenized plant material, but not to AHLs (Ferluga and Venturi, 2009; González et al., 2013). With the exception of BRIP39023, all other Clade II Poaceae isolates have a gene encoding an orphan LuxR family protein that is not found in any other *P. syringae* genome (see Table S2). This gives room for speculation that this protein could be involved in the detection of plant signals that are associated with Poaceae. However, unlike other such sensor proteins, it is not encoded adjacent to a proline immunopeptidase *pip* (Venturi and Fuqua, 2013). The protein shares 78% sequence similarity with the syringomycin biosynthesis regulator, SyrG, of *Psy* B301D (Lu et al., 2002). Syringomycin production is known to be enhanced by p-arbutin (Mo and Gross, 1991), which is a glycosylated hydroquinone found in plant tissues and which also appears to be relatively abundant in wheat (Deisinger et al., 1996). It is not known whether SyrG

actually binds this compound; however, it seems possible that these proteins are able to interact with plant-borne phenolic compounds.

Additionally, the only identified unique gene that is shared among all nine genomes belongs to the LysE/RhtB family of efflux proteins, which are hypothesized to be involved in the export of AHLs (Zakataeva et al., 2006). One proposed function for this family of exporters is protection from interference by AHLs produced from other bacterial species. Interestingly, two recent studies showed the production of AHL-mimicking substances in rice and *Medicago truncatula* (Degrassi et al., 2007; Gao et al., 2003), and therefore, it is possible that such a protein could be involved in the protection of the bacteria from these compounds.

7.3.4.3. Exopolysaccharides

P. syringae strains are known to produce two types of EPS: levan and alginate (Osman et al., 1986). While levan is a branched β-polyfructan (Osman et al., 1986), alginate consists of blocks of non-repeating α-L-glucuronate and β-D-mannuronate, with the latter residue being randomly acetylated (Franklin et al., 2011). Levan is produced from sucrose by the action of a single enzyme, called levansucrase (Li and Ullrich, 2001). *P. syringae* genomes usually contain more than one levansucrase gene located at different genomic sites (Li and Ullrich, 2001). In contrast, alginate biosynthesis is a multi-step process, which involves several enzymes (Franklin et al., 2011). Respective genes are organized in a gene cluster and have an identical architecture in both *P. syringae* and *P. aeruginosa* (Penaloza-Vazquez et al., 1997). One of the major roles of EPS is thought to be protection from desiccation. Furthermore, while alginate was shown to be involved in epiphytic fitness and osmotolerance (Penaloza-Vazquez et al., 1997; Yu et al., 1999), levan is presumed to be a storage polymer (Laue et al., 2006). In addition, most *P. syringae* strains have homologs of genes from *P. aeruginosa* involved in the biosynthesis of another EPS called Psl. Psl, a complex polysaccharide with repeating pentamer units of D-glucose, D-mannose and L-rhamnose (Byrd et al., 2009), was demonstrated to be required for biofilm formation in *P. aeruginosa* (Jackson et al., 2004).

When compared with one another, all Poaceae isolates possess the alginate biosynthesis cluster (*algD-algA*), as well as all other genes involved in its biosynthesis and regulation (*algP*, *algQ*, *algR*, *algZ*, *algC*, *algB*, *kinB*, *algH*, *algU*, *mucA*, *mucB* and *mucD*). With respect to levan production, all strains were identified to have the *lscA* gene, which, however, was shown not to contribute to levan

production (Li and Ullrich, 2001). With the exception of *Por*, all other genomes contained a copy of the *lscC* gene. In *Paf*, both *lscA* and *lscC* are located on two separate contigs, and thus, it is not possible to tell whether the two genes are completely intact. None of the strains had a plasmid-born *lscB* gene. Thus, among the analyzed strains, only *Por* appears to be levan-negative. Last, with respect to *psl* gene content, *P. syringae* is slightly different from *P. aeruginosa* and lacks a homolog of the acyltransferase *pslL*. In addition, the *pslC* gene is located outside of the main gene cluster. However, all nine genomes contain an additional gene between the *pslJ* and *pslK* homologs, which encodes a putative maltose O-acyltransferase family protein. Since the *psl* gene cluster was shown to be transcriptionally active (Hockett et al., 2013), it is possible that *P. syringae* is producing a slightly different version of this EPS. The analyzed *Pcal* strains appear to contain genes for the biosynthesis of both alginate and Psl. A single levansucrase-coding gene is present in all of these strains, with the exception of ES4326, which did not give any hits in our BLAST (Basic Local Alignment Search Tool) search.

7.3.4.4. Type VI secretion system

The Type VI secretion system (T6SS) is a multipurpose protein delivery machinery encoded as a single gene cluster (Shrivastava and Mande, 2008). It is found in the genomes of Gram-negative bacteria, both pathogenic and non-pathogenic, with the highest abundance among Proteobacteria species (Kapitein and Mogk, 2013; Shrivastava and Mande, 2008). It is not uncommon that a single genome contains more than one T6SS gene cluster (Barret et al., 2011; Kapitein and Mogk, 2013). However, different T6SS are found within one genome function independently, and there is no overlap in terms of the translocated effector repertoire (Burtnick et al., 2011; Hachani et al., 2011). In several species, the T6SS was identified to be a pathogenicity factor or at least a virulence enhancer. However, its function is not limited to that, and it was also shown to be involved in inter-species competition, protection from predation and biofilm formation (Kapitein and Mogk, 2013). The first identified effectors for this secretion system were Hcp and VgrG family proteins, which are also structural components of the injection apparatus and are encoded within and outside the T6SS gene cluster (Kapitein and Mogk, 2013). A recent study on *Pto* DC3000 T6SS gene clusters (or HSI, Hcp secretion islands) showed that the Hcp-2 protein encoded in the HSI-2 gene cluster plays a role in competition with other bacterial species, but not in virulence (Haapalainen et al., 2012).

All nine genomes of Poaceae-colonizing strains contain two individual T6SS gene clusters, which belong to the Group 1 and Group 4B loci based on the homology of the respective ClpV proteins, as

classified by Barret and colleagues (Barret et al., 2011). However, the *tssB* gene in the T6SS-2 gene cluster (a Group 4B member) of *Pja* contains a potential frameshift. Both types of T6SS are found in most other sequenced *P. syringae* strains (Barret et al., 2011) and, at least in the case of Poaceae isolates, are located within the same genomic regions. Thus, it appears that these T6SS gene clusters were acquired long ago. A detailed analysis of the T6SS gene clusters from plant-pathogenic bacteria has been performed by Sarris and colleagues (Sarris et al., 2010, 2012), which also depicts the phylogenetic position of both T6SS-1 and T6SS-2 from *Por* and, thus, of their respective homologs from the other analyzed Poaceae isolates. It is notable that none of the genomes encodes a homolog of the above-mentioned Hcp-2 from *Pto* DC3000.

While the HSIs remain conserved, the number of putative orphan effectors varies between the strains. For example, the genome of *Psy* SM encodes five orphan Hcp and three orphan VgrG proteins; *Psy* B64 has only two such Hcp and three such VgrG-coding genes; whereas the genome of *Ppa* encodes for three members of the Hcp family and five VgrG family members located outside the HSIs. The orphan VgrG family proteins sometimes carry an additional domain at their C-terminus and are referred to as "evolved" VgrG proteins. Some of these domains have been described to have a catalytic function, such as actin crosslinking (Kapitein and Mogk, 2013; Pukatzki et al., 2007). Among the analyzed genomes, two VgrG proteins could potentially belong to this group: VgrG-3 of *Psy* SM (PssSM_4133) has an additional domain of unknown function (COG4253), while the C-terminus of VgrG-4 (PssSM_4495) has a weak homology to the anti-sigma factor FecI-like domain. Both proteins are well conserved among *P. syringae* isolates, and in addition, the first one is also found in *P. putida* and *X. oryzae*, while the second one is present in some *P. aeruginosa* strains.

Another group of known T6SS effector encompasses the so-called Type VI lipase effectors (Tle), which are often encoded downstream of VgrG family proteins (Russell et al., 2013). Among the analyzed genomes, there are six that contain genes homologous to PSPTO_5055, which is a member of the Tle3 family: A988_02933 from BRIP39023, POR16_12266 from *Por*, PssB64_01780 from *Psy* B64, PSYJA_00260/PSYJA_00265 from *Pja*, as well as unannotated genes located on contig 283 of the *Paf* genome (accession no. AWUI01000283) and on contig 6 of the *Ppa* genome (accession no. ALAC01000006). All three genes have an overlapping VgrG coding sequence upstream, as well as another ORF (Open Reading Frame) immediately downstream, which could be an immunity protein (based on the architecture of the other Tle3 gene clusters) (Russell et al., 2013).

Notably, A988_02933 contains the conventional GxSxG motif described for the family (residues 237–241), while in all other proteins, the motif is AxSxG (residues 238–242). In addition to that, the genomes of *Por* and BRIP39023 contain putative members of the Tle5 family, POR16_25160 and A988_18422, which are, however, only ~67% homologous to the described family member, PLA107_28855, from *P. syringae* pv. *lachrymans* M301315 (Russell et al., 2013). Nevertheless, both proteins are found within the same genomic region and have a VgrG family protein encoded nearby (POR16_25135 and A988_18442, respectively). Moreover, both of them have a perfect dual HxKxxxxD motif characteristic of phospholipase D and Tle5 family members. In addition, the genome of *Ppa* contains a gene coding for a less related putative Tle5 family protein, located adjacent to one of the T6SS gene clusters (contig 27, accession no. ALAC01000027). Tle family proteins were recently shown to have antibacterial activity, and thus, they could be used by these strains to compete with epiphytic bacteria. No homologs of Tse1, Tse2 or Tse3 from *P. aeruginosa* (Hood et al., 2010) were detected in any of the Poaceae isolates.

7.3.5. Mobile and Extrachromosomal Elements

P. syringae genomes contain a relatively large number of mobile genetic elements (MGE) belonging to various families. The specific composition and the number of MGEs tend to differ from strain to strain (Joardar et al., 2005). More importantly, MGEs are often associated with pathogenicity islands, thus allowing their spread from one cell to another (Lindeberg et al., 2008). It is also known that strains have overcome race-specific resistance of their hosts after the respective avirulence gene was inactivated by an MGE insertion (Greenberg and Vinatzer, 2003). The heterogeneity in MGE composition is also observed for the Poaceae isolates: for example, the genome of *Psy* SM contains at least 21 transposases belonging to nine different families: ISPsy9, ISPpu10, ISPsy6, ISPpu14, ISPsy1, ISRso10, ISPsy5, ISPsy24 and IS200. In comparison, the *Psy* B64 genome encodes at least 22 transposases. The majority of these belong to the ISPsy9 and ISPpu10 families, whereas the remaining ones, with the exception of IS200, represent families not found in *Psy* SM: ISPsy7, ISPsy26, ISPsy5 and ISPsy22. In addition, the genomes encode a number of phage-type integrases/recombinases, the majority of which are truncated. A large portion of the transposable elements have at least one of their ORFs truncated, thus rendering them inactive. However, a number of transposons are present as partial sequences found on contig ends, which makes it impossible to judge their integrity. Moreover, the presence of sequence gaps only allows giving an estimate of their copy number.

Aside from simple insertional elements, the genomes also contain several prophages of variable size. While some of the prophages are very small and contain about 20 ORFs (such as prophage PSSSM-04, PssSM_4692–4712), others are over 57 kb-long (for example, prophage PSSSM-02, PssSM_2181–2256). Interestingly, prophage PSSSM-04 is the only prophage that is common to all Poaceae isolates. Prophages often serve as preferred sites of integration for other MGEs, as well as for genes acquired by horizontal transfer. One such example is prophage PSSB64-02, which has an additional ORF (PssB64_01273) present in comparison to its homolog, PSSSM-04. A more extreme example is prophage PSSSM-03, which contains a number of regions with a differing G + C profile, which is a signature of horizontal gene transfer (PssSM_4287–4288, 4292–4296, 4304–4308, 4312–4313 and 4316–4318). None of the prophages seems to be complete. However, since some components could be shared between different prophages, it is not possible to exclude that at least one is capable of forming functional viral particles. It is noteworthy that the prophage PSSSM-02 contains *recT* and *recE* genes (PssSM_2197 and PssSM_2196, respectively), which encode homologs of the lambda Red Exo/Beta proteins. The RecTE system has been successfully deployed for the recombineering of linear DNA fragments into the *P. syringae* genome (Swingle et al., 2010).

Another interesting type of mobile element present in bacteria are the so-called integrative and conjugative elements (ICEs). These elements are always localized within a set of specific sites within a genome (the so-called *att* sites) and are flanked by direct repeats. They contain genes required for integration and excision and are presumably self-transmittable (Juhas et al., 2009). One of the first discovered ICEs was the *clc* element from *Pseudomonas sp*. B13, which encodes enzymes required for 3-chlorobenzoate degradation (Gaillard et al., 2006). The ICEs were demonstrated to serve as vectors for the spreading of virulence factors, drug and metal resistance genes and other beneficial metabolic traits (Rodríguez-Blanco et al., 2012; Wozniak and Waldor, 2010).

Several examples of ICEs were described for *P. syringae*, as well. The first one characterized was PPHGI-1 from *P. syringae* pv. *phaseolicola* 1302A (Pitman et al., 2005). This genomic region contains a gene encoding the Type III effector, *avrPhpB* (*hopAR1*), and was shown to excise itself upon infiltration into a resistant cultivar of bean. In addition to being located between two tRNA genes and having a recombinase gene required for its excision, it also contains a gene cluster for the biogenesis of conjugative pili. Interestingly, it also contains a copy of the two syringomycin biosynthesis regulator genes, *salA* and *syrF* (*salC*). A similar ICE was also identified in the genome

of *Psy* B728a, where a part of it was replaced by an arsenic and copper resistance gene cluster, and in *Pto* DC3000, where only a part of the ICE is present (Pitman et al., 2005). A similar ICE was recently described for a number of *P. syringae* pv. *actinidiae* strains (McCann et al., 2013).

The genomes of all nine Poaceae isolates have two well-conserved *att* sites, each within a distinct tRNA-Lys gene. However, only seven of the genomes contain an integrated ICE (Table S3A). The exceptions are *Psy* B64 and *Paf*, where the genomic island likely has excised itself at some point after they diverged from *Pja*. Interestingly, several different types of ICEs were identified, which are located in either of the two potential integration sites: in *Ppa* and *Psy* SM, the ICE was adjacent to the *queC* gene, while in all other strains, it was located next to the *clpB* gene. The second observation was that most of the shared Poaceae isolate-specific genes were found adjacent to both *att* sequences, but not associated with the ICE. In general, among the different groups of ICEs, all structural genes were highly conserved, and the observed differences were mainly found within a variable middle part between the *pil* gene cluster and the topoisomerase III gene. The majority of the variable regions encoded heavy metal resistance genes: the ICEs of BRIP34876 and BRIP34881 contain mercury and arsenate resistance operons; *Pja* and *Ppa* have copper and arsenate resistance gene clusters, while in the case of *Psy* SM, chromate, nickel/cobalt and bacteriocin/lantibiotic efflux transporters were found. The ICE of *Por* did not contain any heavy metal resistance genes and was rather similar to the one found in pv. *actinidiae* strain NZ V-13 (McCann et al., 2013). However, it must be noted that the respective *Por* scaffold has a number of gaps and appears to be rather misassembled, and thus, it is possible that some genes were missed. The ICE of BRIP39023, in contrast to the other six ICEs, is relatively similar to PPHGI-1. However, it contains a number of regions with no homology to PPHGI-1, namely a unique region with a number of sensor histidine kinases and transcriptional regulators (A988_21582–A988_21632), as well as another locus surrounding a poly(β-D-mannuronate)-C5-epimerase (A988_21477). Notably, the first locus encodes the T3E HopBF1 (A988_21602), while the second locus contains a truncated copy of another T3E, HopBA1-2 (A988_21482). In addition, it appears that the levansucrase C locus, which is normally located downstream of the respective *att* site, has jumped into the *syrF* homolog, resulting in its duplication. From this comparison, it appears that either there is a "population" of ICEs with distinct variable regions circulating among *P. syringae* strains, or this region is a recombination hotspot, which would also account for the observed diversity.

The last type of mobile genetic elements discussed here are plasmids. *P. syringae* strains often have one or more plasmid, some of which also encode T3Es (Buell et al., 2003; Joardar et al., 2005; Zhao et al., 2005), or other virulence factors, such as coronatine (Bender et al., 1996; Qi et al., 2011) or phytohormone biosynthesis enzymes (Pérez-Martínez et al., 2008). Of the nine Poaceae isolates, four were identified to contain a plasmid based on the presence of a replicase gene adjacent to UV-resistance genes *rulA*/*rulB* (Table S3B). All four identified plasmids belong to the pPT32A family, which is frequently found in *P. syringae* strains (Pérez-Martínez et al., 2008; Zhao et al., 2005). In addition, *Psy* SM appears to have remnants of an integrated plasmid in its genome, which encodes a non-homologous replicase with a 78% similarity to the RepA protein of pPT14-32 from *P. syringae* PT14.

The identified plasmids have a complete or nearly complete Type IV-A secretion system (*virB1*–*virB11* and *virD4* genes (Juhas et al., 2008)). No homologs of *tra*/*trb* genes from pDC3000B (accession AE016854) or *mob* genes were detected. Type IV secretion systems are frequently found on *P. syringae* plasmids; however, whether they play any role in virulence or are only involved in the spreading of the plasmids is yet to be determined (Zhao et al., 2005). Based on the current annotation, most of the proteins encoded on the detected plasmids are hypothetical, and no virulence genes were detected.

7.3.6. Other notable genome components

7.3.6.1. Defense mechanisms against foreign DNA

The CRISPR/Cas (clustered regularly interspaced short palindromic repeats/CRISPR-associated genes) system provides a wide-spread defense mechanism against bacteriophages. CRISPR repeat sequences contain fragments of phage DNA (the so-called spacers), which, upon transcription and processing, are involved in guiding Cas proteins, which, in turn, recognize and cleave the genome of an invading phage (Makarova et al., 2013). The system is adaptive, and new spacers can be acquired upon encountering novel invaders (Toussaint et al., 2012). Among the genomes of Poaceae-colonizing strains, only *Psy* SM was identified to have two putative CRISPRs with more than one spacer sequence (CRISPR 1: 1,828,342–1,828,605; CRISPR 2: 1,828,678–1,829,372; both are inside the *inaZ* gene). However, no putative Cas proteins, which are expected to be encoded in the vicinity (Haft et al., 2005), were detected. It should be noted that *Por* and *Pja* were excluded from the search, due to a high level of genome fragmentation, which would not allow the detection of

direct repeats. In addition, no PIWI-domain proteins, the analogs of the RNA interference complex protein, Argonaute (Makarova et al., 2009), were identified in any of the genomes.

7.3.6.2. Bacteriocins

Bacteriocins are ribosomally synthesized peptides, which might or might not be modified, that show bactericidal activity, usually against a narrow range of species. The producing bacterial species is unaffected by its own bacteriocins, either due to the presence of an immunity protein or by other mechanisms (Cotter et al., 2013). The most common targets of bacteriocins are the cell envelope (Hassan et al., 2012), but they can also interfere with DNA, RNA and protein metabolism (Cotter et al., 2013). Among the target genomes, several bacteriocin-coding genes were identified (Table 2). From those, the product of only one gene is currently characterized. This bacteriocin is called syringacin M and was first described in *Pto* DC3000. It was shown to induce the death of susceptible bacteria by the inhibition of peptidoglycan biosynthesis through degradation of lipid II (Grinter et al., 2012).

Table 2. Putative bacteriocin genes.

Strain	Bacteriocin	Location	Strain	Bacteriocin	Location
BRIP34876	Syringacin M	A979_09507	*Psy* SM	S-type pyocin	PssSM_0863
BRIP34881	Syringacin M	A987_14290		Colicin-DNAse	PssSM_0293 [1]
BRIP39023	S-type pyocin	A988_14544		Bacteriocin	PssSM_4261
	Colicin-DNAse	A988_07204 [1]	*Ppa*	Syringacin M	ALAC01000012
	Cyclized peptide	A988_02638 [2]		S-type pyocin	ALAC01000003
Pja	Syringacin M	PSYJA_15707	*Psy* B64	Syringacin M	PssB64_01273
	Bacteriocin	PSYJA_00994		S-type pyocin	PssB64_04860
	Lasso peptide	PSYJA_17421	*Por*	Colicin-DNAse	POR16_04054 [1]
	Colicin-DNAse	AEAH01000987 [1,3]		Bacteriocin	POR16_27431
Paf	Syringacin M	AWUI01000310		Lasso peptide	DS996947 [4]
	S-type pyocin	AWUI01000345			

[1] Immunity protein-coding gene located nearby. [2] ORF (Open Reading Frame) located inside the specified gene on the opposite strand.[3] The gene contains a potential frameshift. [4] ORF is not annotated, located between POR16_00075 and POR16_00080.

7.3.6.3. Non-Ribosomal Peptide Synthetases/Polyketide Synthetases

Non-ribosomal peptide synthetases (NRPS) and polyketide synthases (PKS) are large multi-modular proteins with enzymatic function found in bacteria and fungi. These proteins are involved in the

biosynthesis of secondary metabolites, such as antibiotics, siderophores, toxins and other bioactive compounds (Nikolouli and Mossialos, 2012). In *P. syringae*, with the exception of phaseolotoxin and tabtoxin, all other phytotoxins, including syringolin, are produced by NRPS/PKS-type enzymes (Arrebola et al., 2012; Bender et al., 1999; Ramel et al., 2009). Other NRPS/PKS gene clusters, which were found in all nine isolates, were genes for the biosynthesis of the siderophore, pyoverdine, and the surfactant, syringofactin. No yersiniabactin biosynthetic cluster was identified in any of the strains. Among NRPS/PKS genes, which were shown to be upregulated *in planta* (Yu et al., 2013), only homologs of Psyr_3722 are found in all strains, whereas a homolog of the PKS gene cluster, Psyr_4311–4315, is exclusively found in *Ppa*. As the products of these genes are unknown, it is currently not possible to draw a conclusion regarding their impact on virulence.

7.4. Experimental section

Genomic sequences used in this study were downloaded from the NCBI (National Center for Biotechnology Information) FTP (File Transfer Protocol) server (ftp://ftp.ncbi.nlm.nih.gov/). Accession numbers of the used genomic sequences are listed in Table S1. The GenBank entries for *Ppa* and *Paf* only contain nucleotide sequence data. Therefore, the genomes were re-annotated using the RAST (Rapid Annotation using Subsystem Technology) server (Aziz et al., 2008) and Prokka (Seemann, 2013). In order to avoid any conflicts, locus_tag qualifiers were set to “panici” and “Paf”, respectively. The phylogenetic tree was generated with MEGA (Molecular Evolutionary Genetics Analysis) version 5.2.2 (Tamura et al., 2011) with 1,000 bootstraps using nucleotide sequences of seven conserved genes: *gyrB*, *gap-1*, *fruK*, *pgi*, *rpoD*, *acnB* and *gltA*. Phylogenetic trees of core Type III effectors were also generated in MEGA 5.2.2 using respective protein sequences and the neighbor-joining method with 1,000 bootstraps. Ortholog cluster searches were performed with the Pan-genome analysis pipeline (Zhao et al., 2012) using the MultiParanoid method with the cut off value set to 80%, an e-value of 10^{-9} and, otherwise, default parameters. Visualization of annotated genomic sequences was done using Artemis (Carver et al., 2012). Homology searches on a gene-to-gene basis were performed using tBLASTN and BLASTP (Altschul et al., 1990; Camacho et al., 2009) against the non-redundant protein sequence database (nr) and other custom databases. The Type III effector sequence database was downloaded from the *Pseudomonas syringae* Genome Resources website (http://www.pseudomonas-syringae.org/). The T3E and phytotoxin compositions of *Por* and *Pja* are based on (Baltrus et al., 2011) and (Mucyn et al., 2014), while the corresponding data for the *Pcal* strains were obtained from (Sarris et al., 2013). Additional analysis of the *Pcal*

strains was performed using IMG/ER (Integrated Microbial Genomes/Expert Review) (Markowitz et al., 2009). Protein domain analysis was done using the NCBI Conserved Domain Database (Marchler-Bauer et al., 2011). Bacteriocin prediction was done using BAGEL3 software (de Jong et al., 2010). The CRISPR search was performed using the CRISPRFinder tool (Grissa et al., 2007).

7.5. Conclusions

Comparative genomics is a powerful tool for the discovery of virulence traits and potential host-specific adaptations. Here, we identified that *P. syringae* strains isolated from Poaceae belong to two different phylogenetic lineages, with each group evolving independently. All of these strains possess the canonical Type III secretion system, but in contrast to the majority of *P. syringae* strains, most Poaceae isolates have a relatively small repertoire of T3Es. Two strains from different phylogenetic lineages were identified to have a second T3SS similar to that of *Rhizobia*. Moreover, in line with previous studies, we have shown that strains pathogenic to the same host could have almost no overlap in T3E composition. In addition, several other putative adaptations were identified, such as a lack of quorum sensing and a number of genes that could be involved in the removal of plant anti-microbial compounds. Moreover, several traits important for leaf surface colonization, such as Type VI secretion systems and bacteriocins, are described. Lastly, the strains were compared with a non-specialized pathogen *P. cannabina* pv. *alisalensis*, which is able to grow on both dicot and monocot plants. The analyses provide a foundation for further experimental validation.

8. General Discussion

Chapter 4 demonstrates that heterologous expression of the *plu1881-plu1877* gene cluster in *P. putida* leads to the production of *bona fide* glidobactin A. This implies that *Ph. luminescens* is also capable of producing this proteasome inhibitor. It still remains an open question whether this bacterium actually produces this compound *in vivo* during some stages of nematode or insect body colonization. NRPS-encoding gene clusters are often induced only under specific conditions, making those difficult to study. Our initial attempts to activate transcription of this gene cluster *in vitro* and subsequently detect a product directly in *Pll* cultures under various conditions were unsuccessful. Therefore, it was decided to explore the possibility of using heterologous hosts. In addition to *P. putida* P3, we also used *Psy* SM and *Psy* B301D-R (Δsyl), which were previously used for expression of the *syl* gene cluster (Ramel et al., 2009). However, for an unknown reason, we detected presence of a glidobactin-like compound in *P. putida* cultures, but not in cultures of other *Pseudomonas* strains. The discovery of *bona fide* glidobactin A in *P. putida* was somewhat surprising. The *plu1881-plu1877* gene cluster is not identical to its homolog in K481-B101 (see Introduction, Figure 4), as it lacks a transcriptional regulator (*glbA*), an MbtH-like protein (*glbE*), and a dioxygenase family protein (*glbH*). While the MbtH-like proteins could be shared between different NRPS-coding gene clusters (Felnagle et al., 2010; Lautru et al., 2007) and the lack of a transcriptional regulator gene suggests that it might be located elsewhere, our original hypothesis was that the product of the *glbH* gene is responsible for the hydroxylation of the lysine residue. Because *glbH* was not needed for glidobactin A biosynthesis in *P. putida* carrying *plu1881-1877*, it appears that it rather plays a role in the modification of the fatty acid tail leading to a minor variant not present in *Pll*. However, it cannot be completely excluded, although seems rather unlikely, that the genome of *P. putida* encodes an enzyme which could substitute for a *glbH* homolog.

In contrast to the *plu1881-1877* gene cluster, the syringolin biosynthetic gene cluster of AP16 (PMI03_05099-PMI03_05102) contains all essential biosynthetic gene homologs (*sylB*-*sylE*). There are, however, two additional ORFs located in close proximity of the *sylB* homolog on the same strand. Such a gene organization might suggest that these genes constitute an operon, in particular because PMI03_05097 encodes a putative NRPS-associated type II thioestherase. However, when tested by RT-PCR, only PMI03_05097 transcripts were detected under syringolin production-inducing conditions. Moreover, the intergenic region between PMI03_5099 (*sylB* homolog) and

PMI03_05097 was not amplifiable by RT-PCR, implying that these genes are transcribed independently. The deletion mutant lacking PMI03_05097 was later found to produce about half the amount of syringolin A as compared to the wild type strain, suggesting that this protein increases the biosynthesis efficiency of this compound. Another special feature of the syrbactin synthetase gene cluster in AP16 is duplicated thiolation domain of the PMI03_05100 (*sylC* homolog), something which so far has not been described for NRPS. In PKS, however, this is rather common, and it has been described that either one or both can be functional. In the latter case this results in increased biosynthesis yield. Here we found that only one thiolation domain (PCP-1) is sufficient for wild type level production of syringolin. It was not possible to ascertain that PCP-2 is non-functional, as our internal sequence deletion might have introduced a detrimental change to the 3D structure of the protein.

The absence of a *sylA* gene homolog, which in *P. syringae* encodes a transcriptional regulator of the *syl* gene cluster, suggests that in AP16 the regulation of expression is somewhat different. Interestingly, however, conditions inducing expression of this gene cluster are the same in *P. syringae* species. The genome of AP16 lacks any putative homologs of both SylA and SalA, which were demonstrated to control syringolin production in *P. syringae* (Ramel et al., 2012). Nonetheless, both species live inside plant tissues, and therefore should be able to sense and respond to plant-borne signals. The SRM_{AFY} medium which we used in this study contains p-arbutin, which is a phenolic compound commonly found in plant tissues (Mo and Gross, 1991), and might be the inducer for this gene cluster. Thus, it is likely that in AP16, the gene cluster is co-regulated together with other plant-born signal-inducible genes, but via a different regulatory circuit.

With regard to a potential role of syringolin in root colonization by AP16, it was not yet possible to set up a proper model using the model plant *Arabidopsis* or local *Populus sp*. When co-incubated with either *Arabidopsis* seedlings or isolated poplar roots on solid MS medium, it was not possible to detect the wild type AP16 tagged with fluorescent protein mCherry inside plant tissue. Interactions between rhizobia and their hosts are usually rather specific (Long, 1996). Thus, it is possible that the bacterium is also specialized and is only able to colonize roots of some particular poplar species.

Chapter 6 provides evidence that syringolin is contributing to virulence and symptom development of *P. syringae* on wheat, as was observed in the case of BRIP34876. Interestingly, this occurs without an effect on intra-foliar bacterial density. However, this effect appears to be dosage

dependent, and over-production of this compound actually limits bacterial growth. The reason to this is unknown, and we could only speculate on why this is the case. What was also unexpected was that there was no difference in entry efficacy between the wild type BRIP34876 and its *sylC*_KO mutant. Thus, it is possible that syringolin has little or no effect on wheat stomata.

In addition to the effects of syringolin, we studied several other virulence traits, including the T3SS. The small number of T3Es is common to all currently sequenced group II Poaceae isolates, yet it was surprising that knocking out the T3SS made these strains completely avirulent. Moreover, the effects of deleting individual non-core effectors appeared to have a minor effect on endophytic growth of *Psy* SM. However, the degree by which individual mutations (Δ*hopBA1*, Δ*hopA2*, or Δ*hopAZ1*) affected the endophytic populations was different, and the largest decrease was observed for the Δ*hopBA1*strain. The T3E HopBA1 is common to all phylogenetic clade II Poaceae isolates, as identified in chapter 7, hence, it is logical to suggest that this protein plays an important role and its deletion has more pronounced consequences. In addition, it was interesting to find that mutation of the *pslD* gene resulted in a small but consistent reduction of endophytic growth of *Psy* SM. Therefore a future study needs to be directed towards clarifying whether this effect is truly due to deficiency in Psl production, or rather is due to an indirect effect. If Psl is indeed produced and plays a role in virulence, then it would be also of value to verify the structure of this EPS in *P. syringae*, as well as how its production is regulated.

Last but not least, our genomics-based comparison highlighted several genes conserved among Poaceae isolates. These candidates might provide us with more insights into the biology of interactions between Poaceae family members and their pathogens. Therefore, a mutational analysis of the identified genes needs to be performed in the future. Moreover, it would be interesting to determine targets of individual T3Es, in particular of HopBA1 and HopM1. This might point out important players and hubs of monocot defense signaling. Another point could be to study the significance of other phytotoxins, such as mangotoxin or syringomycin. In addition, it would also be interesting to experimentally validate other traits identified by this analysis, such as putative type VI effectors or the functionality of ICEs.

9. Literature

Akiyoshi, D.E., Regier, D.A., and Gordon, M.P. (1987). Cytokinin production by *Agrobacterium* and *Pseudomonas* spp. J. Bacteriol. *169*, 4242–4248.

Alfano, J.R., Charkowski, A.O., Deng, W.L., Badel, J.L., Petnicki-Ocwieja, T., van Dijk, K., and Collmer, A. (2000). The *Pseudomonas syringae* Hrp pathogenicity island has a tripartite mosaic structure composed of a cluster of type III secretion genes bounded by exchangeable effector and conserved effector loci that contribute to parasitic fitness and pathogenicit. Proc. Natl. Acad. Sci. U. S. A. *97*, 4856–4861.

Almeida, N.F., Yan, S., Lindeberg, M., Studholme, D.J., Schneider, D.J., Condon, B., Liu, H., Viana, C.J., Warren, A., Evans, C., et al. (2009). A draft genome sequence of *Pseudomonas syringae* pv. *tomato* T1 reveals a type III effector repertoire significantly divergent from that of *Pseudomonas syringae* pv. *tomato* DC3000. Mol. Plant. Microbe. Interact. *22*, 52–62.

Altschul, S.F., Gish, W., Miller, W., Myers, E.W., and Lipman, D.J. (1990). Basic local alignment search tool. J. Mol. Biol. *215*, 403–410.

Amrein, H., Makart, S., Granado, J., Shakya, R., Schneider-Pokorny, J., and Dudler, R. (2004). Functional analysis of genes involved in the synthesis of syringolin A by *Pseudomonas syringae* pv. *syringae* B301 D-R. Mol. Plant. Microbe. Interact. *17*, 90–97.

Ansari, M.Z., Yadav, G., Gokhale, R.S., and Mohanty, D. (2004). NRPS-PKS: a knowledge-based resource for analysis of NRPS/PKS megasynthases. Nucleic Acids Res. *32*, W405–13.

Archer, C.R., Koomoa, D.-L.T., Mitsunaga, E.M., Clerc, J., Shimizu, M., Kaiser, M., Schellenberg, B., Dudler, R., and Bachmann, A.S. (2010). Syrbactin class proteasome inhibitor-induced apoptosis and autophagy occurs in association with p53 accumulation and Akt/PKB activation in neuroblastoma. Biochem. Pharmacol. *80*, 170–178.

Arnold, D.L., Lovell, H.C., Jackson, R.W., and Mansfield, J.W. (2011). *Pseudomonas syringae* pv. *phaseolicola*: from "has bean" to supermodel. Mol. Plant Pathol. *12*, 617–627.

Arrebola, E., Cazorla, F.M., Codina, J.C., Gutiérrez-Barranquero, J.A., Pérez-García, A., and de Vicente, A. (2009). Contribution of mangotoxin to the virulence and epiphytic fitness of *Pseudomonas syringae* pv. *syringae*. Int. Microbiol. *12*, 87–95.

Arrebola, E., Cazorla, F.M., Perez-García, A., and de Vicente, A. (2011). Chemical and metabolic aspects of antimetabolite toxins produced by *Pseudomonas syringae* pathovars. Toxins (Basel). *3*, 1089–1110.

Arrebola, E., Carrión, V.J., Cazorla, F.M., Pérez-García, A., Murillo, J., and de Vicente, A. (2012). Characterisation of the mgo operon in Pseudomonas syringae pv. syringae UMAF0158 that is required for mangotoxin production. BMC Microbiol. *12*, 10.

Artsimovitch, I., Svetlov, V., Nemetski, S.M., Epshtein, V., Cardozo, T., and Nudler, E. (2011). Tagetitoxin inhibits RNA polymerase through trapping of the trigger loop. J. Biol. Chem. *286*, 40395–40400.

Ashburner, M., Ball, C.A., Blake, J.A., Botstein, D., Butler, H., Cherry, J.M., Davis, A.P., Dolinski, K., Dwight, S.S., Eppig, J.T., et al. (2000). Gene ontology: tool for the unification of biology. Nat. Genet. *25*, 25–29.

Ausubel, F.M., Brent, R., Kingston, R.E., Moore, D.D., Seidman, J.G., Smith, J.A., Struhl, K., and (eds.) (1998). Current Protocols in Molecular Biology (New York, USA: John Wiley & Sons).

Aziz, R.K., Bartels, D., Best, A.A., DeJongh, M., Disz, T., Edwards, R.A., Formsma, K., Gerdes, S., Glass, E.M., Kubal, M., et al. (2008). The RAST Server: rapid annotations using subsystems technology. BMC Genomics *9*, 75.

Badel, J.L., Shimizu, R., Oh, H.-S., and Collmer, A. (2006). A *Pseudomonas syringae* pv. *tomato avrE1*/*hopM1* mutant is severely reduced in growth and lesion formation in tomato. Mol. Plant-Microbe Interact. *19*, 99–111.

Baltrus, D.A., Nishimura, M.T., Romanchuk, A., Chang, J.H., Mukhtar, M.S., Cherkis, K., Roach, J., Grant, S.R., Jones, C.D., and Dangl, J.L. (2011). Dynamic evolution of pathogenicity revealed by sequencing and comparative genomics of 19 *Pseudomonas syringae* isolates. PLoS Pathog. *7*, e1002132.

Baltrus, D.A., Dougherty, K., Beckstrom-Sternberg, S.M., Beckstrom-Sternberg, J.S., and Foster, J.T. (2013). Incongruence between multi-locus sequence analysis (MLSA) and whole-genome-based phylogenies: Pseudomonas syringae pathovar pisi as a cautionary tale. Mol. Plant Pathol. In press.

Barret, M., Egan, F., Fargier, E., Morrissey, J.P., and O'Gara, F. (2011). Genomic analysis of the type VI secretion systems in Pseudomonas spp.: novel clusters and putative effectors uncovered. Microbiology *157*, 1726–1739.

Bender, C., Palmer, D., Peñaloza-Vázquez, A., Rangaswamy, V., and Ullrich, M. (1996). Biosynthesis of coronatine, a thermoregulated phytotoxin produced by the phytopathogen *Pseudomonas syringae*. Arch. Microbiol. *166*, 71–75.

Bender, C.L., Alarcón-Chaidez, F., and Gross, D.C. (1999). *Pseudomonas syringae* phytotoxins: mode of action, regulation, and biosynthesis by peptide and polyketide synthetases. Microbiol. Mol. Biol. Rev. MMBR *63*, 266–292.

Bernard, C.S., Brunet, Y.R., Gueguen, E., and Cascales, E. (2010). Nooks and crannies in type VI secretion regulation. J. Bacteriol. *192*, 3850–3860.

Bertani, G. (1951). Studies on lysogenesis. I. The mode of phage liberation by lysogenic *Escherichia coli*. J. Bacteriol. *62*, 293–300.

Bhattacharjee, S., Halane, M.K., Kim, S.H., and Gassmann, W. (2011). Pathogen effectors target *Arabidopsis* EDS1 and alter its interactions with immune regulators. Science *334*, 1405–1408.

Von Bodman, S.B., Bauer, W.D., and Coplin, D.L. (2003). Quorum sensing in plant-pathogenic bacteria. Annu. Rev. Phytopathol. *41*, 455–482.

Boller, T., and Felix, G. (2009). A renaissance of elicitors: perception of microbe-associated molecular patterns and danger signals by pattern-recognition receptors. Annu. Rev. Plant Biol. *60*, 379–406.

Bordiec, S., Paquis, S., Lacroix, H., Dhondt, S., Ait Barka, E., Kauffmann, S., Jeandet, P., Mazeyrat-Gourbeyre, F., Clément, C., Baillieul, F., et al. (2011). Comparative analysis of defence responses induced by the endophytic plant growth-promoting rhizobacterium *Burkholderia phytofirmans* strain PsJN and the non-host bacterium *Pseudomonas syringae* pv. *pisi* in grapevine cell suspensions. J. Exp. Bot. *62*, 595–603.

Bradbury, J.F. (1986). *Pseudomonas syringae* pv. *syringae*. In Guide to Plant Pathogenic Bacteria, (Kew, England: CAB International Mycological Institute), pp. 175–177.

Brencic, A., and Winans, S.C. (2005). Detection of and response to signals involved in host-microbe interactions by plant-associated bacteria. Microbiol. Mol. Biol. Rev. *69*, 155–194.

Van den Broek, D., Bloemberg, G. V, and Lugtenberg, B. (2005). The role of phenotypic variation in rhizosphere *Pseudomonas* bacteria. Environ. Microbiol. *7*, 1686–1697.

Brown, S.D., Utturkar, S.M., Klingeman, D.M., Johnson, C.M., Martin, S.L., Land, M.L., Lu, T.-Y.S., Schadt, C.W., Doktycz, M.J., and Pelletier, D.A. (2012). Twenty-one genome sequences from *Pseudomonas* species and 19 genome sequences from diverse bacteria isolated from the rhizosphere and endosphere of *Populus deltoides*. J. Bacteriol. *194*, 5991–5993.

Buell, C.R., Joardar, V., Lindeberg, M., Selengut, J., Paulsen, I.T., Gwinn, M.L., Dodson, R.J., Deboy, R.T., Durkin, A.S., Kolonay, J.F., et al. (2003). The complete genome sequence of the *Arabidopsis* and tomato pathogen *Pseudomonas syringae* pv. *tomato* DC3000. Proc. Natl. Acad. Sci. U. S. A. *100*, 10181–10186.

Bullock, W.O., Fernandez, J.M., and Short, J.M. (1987). XL1-Blue: a high efficiency plasmid transforming *recA Escherichia coli* strain with beta-galactosidase selection. Biotechniques *5*, 376–378.

Burch, A.Y., Shimada, B.K., Mullin, S.W.A., Dunlap, C.A., Bowman, M.J., and Lindow, S.E. (2012). Pseudomonas syringae coordinates production of a motility-enabling surfactant with flagellar assembly. J. Bacteriol. *194*, 1287–1298.

Burtnick, M.N., Brett, P.J., Harding, S. V, Ngugi, S.A., Ribot, W.J., Chantratita, N., Scorpio, A., Milne, T.S., Dean, R.E., Fritz, D.L., et al. (2011). The cluster 1 type VI secretion system is a major virulence determinant in *Burkholderia pseudomallei*. Infect. Immun. *79*, 1512–1525.

Büttner, D. (2012). Protein export according to schedule: architecture, assembly, and regulation of type III secretion systems from plant- and animal-pathogenic bacteria. Microbiol. Mol. Biol. Rev. *76*, 262–310.

Byrd, M.S., Sadovskaya, I., Vinogradov, E., Lu, H., Sprinkle, A.B., Richardson, S.H., Ma, L., Ralston, B., Parsek, M.R., Anderson, E.M., et al. (2009). Genetic and biochemical analyses of the *Pseudomonas aeruginosa* Psl exopolysaccharide reveal overlapping roles for polysaccharide synthesis enzymes in Psl and LPS production. Mol. Microbiol. *73*, 622–638.

Cai, R., Lewis, J., Yan, S., Liu, H., Clarke, C.R., Campanile, F., Almeida, N.F., Studholme, D.J., Lindeberg, M., Schneider, D., et al. (2011). The plant pathogen *Pseudomonas syringae* pv. *tomato* is genetically monomorphic and under strong selection to evade tomato immunity. PLoS Pathog. *7*, e1002130.

Camacho, C., Coulouris, G., Avagyan, V., Ma, N., Papadopoulos, J., Bealer, K., and Madden, T.L. (2009). BLAST+: architecture and applications. BMC Bioinformatics *10*, 421.

Carrión, V.J., Gutiérrez-Barranquero, J.A., Arrebola, E., Bardaji, L., Codina, J.C., de Vicente, A., Cazorla, F.M., and Murillo, J. (2013). The mangotoxin biosynthetic operon (mbo) is specifically distributed within *Pseudomonas syringae* genomospecies 1 and was acquired only once during evolution. Appl. Environ. Microbiol. *79*, 756–767.

Carver, T., Harris, S.R., Berriman, M., Parkhill, J., and McQuillan, J.A. (2012). Artemis: an integrated platform for visualization and analysis of high-throughput sequence-based experimental data. Bioinformatics *28*, 464–469.

Cazorla, F.M., Torés, J.A., Olalla, L., Pérez-García, A., Farré, J.M., and de Vicente, A. (1998). Bacterial Apical Necrosis of Mango in Southern Spain: A Disease Caused by *Pseudomonas syringae* pv. *syringae*. Phytopathology *88*, 614–620.

Ceri, M., Ortabozkoyun, L., Altay, M., Unverdi, S., Kurultak, I., Huddam, B., Kilic, F., Yilmaz, R., and Duranay, M. (2010). Peritonitis due to *Pseudomonas stutzeri*, an organism that may be difficult to culture. Perit. Dial. Int. *30*, 484–486.

Cesari, S., Thilliez, G., Ribot, C., Chalvon, V., Michel, C., Jauneau, A., Rivas, S., Alaux, L., Kanzaki, H., Okuyama, Y., et al. (2013). The rice resistance protein pair RGA4/RGA5 recognizes the *Magnaporthe oryzae* effectors AVR-Pia and AVR1-CO39 by direct binding. Plant Cell *25*, 1463–1481.

Challis, G.L. (2008). Mining microbial genomes for new natural products and biosynthetic pathways. Microbiology *154*, 1555–1569.

Le Chatelier, E., Nielsen, T., Qin, J., Prifti, E., Hildebrand, F., Falony, G., Almeida, M., Arumugam, M., Batto, J.-M., Kennedy, S., et al. (2013). Richness of human gut microbiome correlates with metabolic markers. Nature *500*, 541–546.

Chatterjee, A., Cui, Y., Yang, H., Collmer, A., Alfano, J.R., and Chatterjee, A.K. (2003). GacA, the response regulator of a two-component system, acts as a master regulator in *Pseudomonas syringae* pv. *tomato* DC3000 by controlling regulatory RNA, transcriptional activators, and alternate sigma factors. Mol. Plant-Microbe Interact. *16*, 1106–1117.

Chin-A-Woeng, T.F., Bloemberg, G. V, Mulders, I.H., Dekkers, L.C., and Lugtenberg, B.J. (2000). Root colonization by phenazine-1-carboxamide-producing bacterium *Pseudomonas chlororaphis* PCL1391 is essential for biocontrol of tomato foot and root rot. Mol. Plant. Microbe. Interact. *13*, 1340–1345.

Chinchilla, D., Bauer, Z., Regenass, M., Boller, T., and Felix, G. (2006). The *Arabidopsis* receptor kinase FLS2 binds flg22 and determines the specificity of flagellin perception. Plant Cell *18*, 465–476.

Christensen, B.B., Sternberg, C., Andersen, J.B., Palmer, R.J., Nielsen, A.T., Givskov, M., and Molin, S. (1999). Molecular tools for study of biofilm physiology. Methods Enzymol. *310*, 20–42.

Cintas, N.A., Koike, S.T., and Bull, C.T. (2002). A New Pathovar, *Pseudomonas syringae* pv. *alisalensis* pv. nov., Proposed for the Causal Agent of Bacterial Blight of Broccoli and Broccoli Raab. Plant Dis. *86*, 992–998.

Clark, D.J., and Maaløe, O. (1967). DNA replication and the division cycle in *Escherichia coli*. J. Mol. Biol. *23*, 99–112.

Clarke, C.R., Cai, R., Studholme, D.J., Guttman, D.S., and Vinatzer, B.A. (2010). Pseudomonas syringae strains naturally lacking the classical *P. syringae* hrp/hrc Locus are common leaf colonizers equipped with an atypical type III secretion system. Mol. Plant. Microbe. Interact. *23*, 198–210.

Cody, Y.S., and Gross, D.C. (1987). Outer membrane protein mediating iron uptake via pyoverdin$_{pss}$, the fluorescent siderophore produced by *Pseudomonas syringae* pv. *syringae*. J. Bacteriol. *169*, 2207–2214.

Conn, V.M., Walker, A.R., and Franco, C.M.M. (2008). Endophytic actinobacteria induce defense pathways in *Arabidopsis thaliana*. Mol. Plant. Microbe. Interact. *21*, 208–218.

Conrath, U., Beckers, G.J.M., Flors, V., García-Agustín, P., Jakab, G., Mauch, F., Newman, M.-A., Pieterse, C.M.J., Poinssot, B., Pozo, M.J., et al. (2006). Priming: getting ready for battle. Mol. Plant. Microbe. Interact. *19*, 1062–1071.

Costacurta, A., and Vanderleyden, J. (2008). Synthesis of Phytohormones by Plant-Associated Bacteria. Crit. Rev. Microbiol. *21*, 1–18.

Cotter, P.D., Ross, R.P., and Hill, C. (2013). Bacteriocins - a viable alternative to antibiotics? Nat. Rev. Microbiol. *11*, 95–105.

Coulthurst, S.J. (2013). The Type VI secretion system – a widespread and versatile cell targeting system. Res. Microbiol. *164*, 640–654.

Crawford, J.M., Kontnik, R., and Clardy, J. (2010). Regulating alternative lifestyles in entomopathogenic bacteria. Curr. Biol. *20*, 69–74.

Cunnac, S., Lindeberg, M., and Collmer, A. (2009). *Pseudomonas syringae* type III secretion system effectors: repertoires in search of functions. Curr. Opin. Microbiol. *12*, 53–60.

Cunnac, S., Chakravarthy, S., Kvitko, B.H., Russell, A.B., Martin, G.B., and Collmer, A. (2011). Genetic disassembly and combinatorial reassembly identify a minimal functional repertoire of type III effectors in *Pseudomonas syringae*. Proc. Natl. Acad. Sci. U. S. A. *108*, 2975–2980.

DebRoy, S., Thilmony, R., Kwack, Y.-B., Nomura, K., and He, S.Y. (2004). A family of conserved bacterial effectors inhibits salicylic acid-mediated basal immunity and promotes disease necrosis in plants. Proc. Natl. Acad. Sci. U. S. A. *101*, 9927–9932.

Degrassi, G., Devescovi, G., Solis, R., Steindler, L., and Venturi, V. (2007). *Oryza sativa* rice plants contain molecules that activate different quorum-sensing N-acyl homoserine lactone biosensors and are sensitive to the specific AiiA lactonase. FEMS Microbiol. Lett. *269*, 213–220.

Deisinger, P.J., Hill, T.S., and English, J.C. (1996). Human exposure to naturally occurring hydroquinone. J. Toxicol. Environ. Health *47*, 31–46.

Delcher, A.L., Bratke, K.A., Powers, E.C., and Salzberg, S.L. (2007). Identifying bacterial genes and endosymbiont DNA with Glimmer. Bioinformatics *23*, 673–679.

Deng, W.L., and Huang, H.C. (1999). Cellular locations of *Pseudomonas syringae* pv. *syringae* HrcC and HrcJ proteins, required for harpin secretion via the type III pathway. J. Bacteriol. *181*, 2298–2301.

Deng, W.-L., Preston, G., Collmer, A., Chang, C.-J., and Huang, H.-C. (1998). Characterization of the *hrpC* and *hrpRS* operons of *Pseudomonas syringae* pathovars syringae, tomato, and glycinea and analysis of the ability of *hrpF*, *hrpG*, *hrcC*, *hrpT*, and *hrpV* mutants to elicit the h. J. Bacteriol. *180*, 4523–4531.

Denison, R.F., and Kiers, E.T. (2011). Life histories of symbiotic rhizobia and mycorrhizal fungi. Curr. Biol. *21*, R775–85.

Denny, T.P. (1988). Phenotypic Diversity in *Pseudomonas syringae* pv. *tomato*. Microbiology *134*, 1939–1948.

Denny, T.P. (1995). Involvement of bacterial polysaccharides in plant pathogenesis. Annu. Rev. Phytopathol. *33*, 173–197.

Denny, T.P., Gilmour, M.N., and Selander, R.K. (1988). Genetic Diversity and Relationships of Two Pathovars of *Pseudomonas syringae*. Microbiology *134*, 1949–1960.

Derksen, H., Rampitsch, C., and Daayf, F. (2013). Signaling cross-talk in plant disease resistance. Plant Sci. *207*, 79–87.

Deslandes, L., and Rivas, S. (2012). Catch me if you can: bacterial effectors and plant targets. Trends Plant Sci. *17*, 644–655.

Dodds, P.N., and Rathjen, J.P. (2010). Plant immunity: towards an integrated view of plant-pathogen interactions. Nat. Rev. Genet. *11*, 539–548.

Driscoll, W.W., Pepper, J.W., Pierson, L.S., and Pierson, E.A. (2011). Spontaneous Gac mutants of *Pseudomonas* biological control strains: cheaters or mutualists? Appl. Environ. Microbiol. *77*, 7227–7235.

Dudler, R. (2013). Manipulation of host proteasomes as a virulence mechanism of plant pathogens. Annu. Rev. Phytopathol. *51*, 521–542.

Dudler, R. (2014). The role of bacterial phytotoxins in inhibiting the eukaryotic proteasome. Trends Microbiol. *22*, 28–35.

Dudnik, A., and Dudler, R. (2013a). Non contiguous-finished genome sequence of *Pseudomonas syringae* pathovar *syringae* strain B64 isolated from wheat. Stand. Genomic Sci. *8*, 420–429.

Dudnik, A., and Dudler, R. (2013b). High-Quality Draft Genome Sequence of *Pseudomonas syringae* pv. *Syringae* Strain SM, Isolated from Wheat. Genome Announc. *1*, e00610–13.

Dudnik, A., and Dudler, R. (2014). Genomics-based exploration of virulence determinants and host-specific adaptations of *Pseudomonas syringae* strains isolated from grasses. Pathogens *3*, 121–148.

Dudnik, A., Bigler, L., and Dudler, R. (2013). Heterologous expression of a Photorhabdus luminescens syrbactin-like gene cluster results in production of the potent proteasome inhibitor glidobactin A. Microbiol. Res. *168*, 73–76.

Dumenyo, C.K., Mukherjee, A., Chun, W., and Chatterjee, A.K. (1998). Genetic and physiological evidence for the production of N-acyl homoserine lactones by *Pseudomonas syringae* pv. *syringae* and other fluorescent plant pathogenic Pseudomonas species. Eur. J. Plant Pathol. *104*, 569–582.

Durrant, W.E., and Dong, X. (2004). Systemic acquired resistance. Annu. Rev. Phytopathol. *42*, 185–209.

Elasri, M., Delorme, S., Lemanceau, P., Stewart, G., Laue, B., Glickmann, E., Oger, P.M., and Dessaux, Y. (2001). Acyl-homoserine lactone production is more common among plant-associated *Pseudomonas* spp. than among soilborne *Pseudomonas* spp. Appl. Environ. Microbiol. *67*, 1198–1209.

El-Sayed, A.K., Hothersall, J., Cooper, S.M., Stephens, E., Simpson, T.J., and Thomas, C.M. (2003). Characterization of the mupirocin biosynthesis gene cluster from Pseudomonas fluorescens NCIMB 10586. Chem. Biol. *10*, 419–430.

Farrand, S.K. (1990). *Agrobacterium radiobacter* strain K84: a model biocontrol system. In New Directions in Biological Control. Alternatives for Suppressing Agricultural Pests and Diseases., R.R. Baker, and P.E. Dunn, eds. (New York, USA: Wiley-Liss Inc.), pp. 679–691.

Feil, H., Feil, W.S., Chain, P., Larimer, F., DiBartolo, G., Copeland, A., Lykidis, A., Trong, S., Nolan, M., Goltsman, E., et al. (2005). Comparison of the complete genome sequences of *Pseudomonas syringae* pv. *syringae* B728a and pv. *tomato* DC3000. Proc. Natl. Acad. Sci. U. S. A. *102*, 11064–11069.

Felnagle, E.A., Barkei, J.J., Park, H., Podevels, A.M., McMahon, M.D., Drott, D.W., and Thomas, M.G. (2010). MbtH-like proteins as integral components of bacterial nonribosomal peptide synthetases. Biochemistry *49*, 8815–8817.

Ferluga, S., and Venturi, V. (2009). OryR is a LuxR-family protein involved in interkingdom signaling between pathogenic *Xanthomonas oryzae* pv. *oryzae* and rice. J. Bacteriol. *191*, 890–897.

Field, D., Garrity, G., Gray, T., Morrison, N., Selengut, J., Sterk, P., Tatusova, T., Thomson, N., Allen, M.J., Angiuoli, S. V, et al. (2008). The minimum information about a genome sequence (MIGS) specification. Nat. Biotechnol. *26*, 541–547.

Finking, R., and Marahiel, M.A. (2004). Biosynthesis of nonribosomal peptides. Annu. Rev. Microbiol. *58*, 453–488.

Franklin, M.J., Nivens, D.E., Weadge, J.T., and Howell, P.L. (2011). Biosynthesis of the Pseudomonas aeruginosa Extracellular Polysaccharides, Alginate, Pel, and Psl. Front. Microbiol. *2*, 167.

Friedman, L., and Kolter, R. (2004). Two genetic loci produce distinct carbohydrate-rich structural components of the *Pseudomonas aeruginosa* biofilm matrix. J. Bacteriol. *186*, 4457–4465.

Gaillard, M., Vallaeys, T., Vorhölter, F.J., Minoia, M., Werlen, C., Sentchilo, V., Pühler, A., and van der Meer, J.R. (2006). The clc element of *Pseudomonas sp*. strain B13, a genomic island with various catabolic properties. J. Bacteriol. *188*, 1999–2013.

Galán, J.E., and Wolf-Watz, H. (2006). Protein delivery into eukaryotic cells by type III secretion machines. Nature *444*, 567–573.

Gangadharan, A., Sreerekha, M.-V., Whitehill, J., Ham, J.H., and Mackey, D. (2013). The *Pseudomonas syringae* pv. *tomato* Type III Effector HopM1 Suppresses *Arabidopsis* Defenses Independent of Suppressing Salicylic Acid Signaling and of Targeting AtMIN7. PLoS One *8*, e82032.

Gao, M., Teplitski, M., Robinson, J.B., and Bauer, W.D. (2003). Production of Substances by *Medicago truncatula* that Affect Bacterial Quorum Sensing. Mol. Plant-Microbe Interact. *16*, 827–834.

Gardan, L., Bollet, C., Abu Ghorrah, M., Grimont, F., and Grimont, P.A.D. (1992). DNA Relatedness among the Pathovar Strains of *Pseudomonas syringae* subsp. *savastanoi* Janse (1982) and Proposal of *Pseudomonas savastanoi* sp. nov. Int. J. Syst. Bacteriol. *42*, 606–612.

Gardan, L., Shafik, H., Belouin, S., Broch, R., Grimont, F., and Grimont, P.A.D. (1999). DNA relatedness among the pathovars of *Pseudomonas syringae* and description of *Pseudomonas tremae* sp. nov. and *Pseudomonas cannabina* sp. nov. (ex Sutic and Dowson 1959). Int. J. Syst. Bacteriol. *49*, 469–478.

Gardiner, D.M., Stiller, J., Covarelli, L., Lindeberg, M., Shivas, R.G., and Manners, J.M. (2013). Genome Sequences of *Pseudomonas* spp. Isolated from Cereal Crops. Genome Announc. *1*, e00209–13.

Garrity, G., Bell, J., and Lilburn, T. (2005). Order IX. Pseudomonales. In Bergey's Manual of Systematic Bacteriology, Second Edition, Volume 2, Part B, G. Garrity, D. Brenner, N. Krieg, and J. Staley, eds. (Springer, New York), p. 323.

Gasson, M.J. (1980). Indicator Technique for Antimetabolic Toxin Production by Phytopathogenic Species of *Pseudomonas*. Appl. Envir. Microbiol. *39*, 25–29.

Gazi, A.D., Sarris, P.F., Fadouloglou, V.E., Charova, S.N., Mathioudakis, N., Panopoulos, N.J., and Kokkinidis, M. (2012). Phylogenetic analysis of a gene cluster encoding an additional, rhizobial-like type III secretion system that is narrowly distributed among *Pseudomonas syringae* strains. BMC Microbiol. *12*, 188.

Geier, G., and Geider, K. (1993). Characterization and influence on virulence of the levansucrase gene from the fireblight pathogen *Erwinia amylovora*. Physiol. Mol. Plant Pathol. *42*, 387–404.

Gellatly, S.L., and Hancock, R.E.W. (2013). *Pseudomonas aeruginosa*: new insights into pathogenesis and host defenses. Pathog. Dis. *67*, 159–173.

Gilbert, S.F., Sapp, J., and Tauber, A.I. (2012). A symbiotic view of life: we have never been individuals. Q. Rev. Biol. *87*, 325–341.

Gimenez-Ibanez, S., and Rathjen, J.P. (2010). The case for the defense: plants versus *Pseudomonas syringae*. Microbes Infect. *12*, 428–437.

Gimenez-Ibanez, S., Hann, D.R., Ntoukakis, V., Petutschnig, E., Lipka, V., and Rathjen, J.P. (2009). AvrPtoB Targets the LysM Receptor Kinase CERK1 to Promote Bacterial Virulence on Plants. Curr. Biol. *19*, 423–429.

Gironde, S., and Manceau, C. (2012). Housekeeping gene sequencing and multilocus variable-number tandem-repeat analysis to identify subpopulations within *Pseudomonas syringae* pv. *maculicola* and *Pseudomonas syringae* pv. *tomato* that correlate with host specificity. Appl. Environ. Microbiol. *78*, 3266–3279.

Glass, M.B., Steigerwalt, A.G., Jordan, J.G., Wilkins, P.P., and Gee, J.E. (2006). *Burkholderia oklahomensis* sp. nov., a *Burkholderia pseudomallei*-like species formerly known as the Oklahoma strain of *Pseudomonas pseudomallei*. Int. J. Syst. Evol. Microbiol. *56*, 2171–2176.

Glickmann, E., Gardan, L., Jacquet, S., Hussain, S., Elasri, M., Petit, A., and Dessaux, Y. (1998). Auxin production is a common feature of most pathovars of *Pseudomonas syringae*. Mol. Plant. Microbe. Interact. *11*, 156–162.

Glover, B., Kachroo, P., Kachroo, A., and Robin, G.P. (2013). Systemic signaling during plant defense. Curr. Opin. Plant Biol. *16*, 527–533.

Göhre, V., Spallek, T., Häweker, H., Mersmann, S., Mentzel, T., Boller, T., de Torres, M., Mansfield, J.W., and Robatzek, S. (2008). Plant pattern-recognition receptor FLS2 is directed for degradation by the bacterial ubiquitin ligase AvrPtoB. Curr. Biol. *18*, 1824–1832.

González, A.J., Landeras, E., and Mendoza, M.C. (2000). Pathovars of *Pseudomonas syringae* causing bacterial brown spot and halo blight in *Phaseolus vulgaris* L. are distinguishable by ribotyping. Appl. Environ. Microbiol. *66*, 850–854.

González, J.F., Myers, M.P., and Venturi, V. (2013). The inter-kingdom solo OryR regulator of *Xanthomonas oryzae* is important for motility. Mol. Plant Pathol. *14*, 211–221.

Goris, J., Konstantinidis, K.T., Klappenbach, J.A., Coenye, T., Vandamme, P., and Tiedje, J.M. (2007). DNA-DNA hybridization values and their relationship to whole-genome sequence similarities. Int. J. Syst. Evol. Microbiol. *57*, 81–91.

Green, S., Studholme, D.J., Laue, B.E., Dorati, F., Lovell, H., Arnold, D., Cottrell, J.E., Bridgett, S., Blaxter, M., Huitema, E., et al. (2010). Comparative genome analysis provides insights into the evolution and adaptation of *Pseudomonas syringae* pv. *aesculi* on *Aesculus hippocastanum*. PLoS One *5*, e10224.

Greenberg, J.T., and Vinatzer, B.A. (2003). Identifying type III effectors of plant pathogens and analyzing their interaction with plant cells. Curr. Opin. Microbiol. *6*, 20–28.

Grinter, R., Roszak, A.W., Cogdell, R.J., Milner, J.J., and Walker, D. (2012). The crystal structure of the lipid II-degrading bacteriocin syringacin M suggests unexpected evolutionary relationships between colicin M-like bacteriocins. J. Biol. Chem. *287*, 38876–38888.

Grissa, I., Vergnaud, G., and Pourcel, C. (2007). CRISPRFinder: a web tool to identify clustered regularly interspaced short palindromic repeats. Nucleic Acids Res. *35*, W52–7.

Groll, M., Schellenberg, B., Bachmann, A.S., Archer, C.R., Huber, R., Powell, T.K., Lindow, S., Kaiser, M., and Dudler, R. (2008). A plant pathogen virulence factor inhibits the eukaryotic proteasome by a novel mechanism. Nature *452*, 755–758.

Gu, L., Eisman, E.B., Dutta, S., Franzmann, T.M., Walter, S., Gerwick, W.H., Skiniotis, G., and Sherman, D.H. (2011). Tandem acyl carrier proteins in the curacin biosynthetic pathway promote consecutive multienzyme reactions with a synergistic effect. Angew. Chem. Int. Ed. Engl. *50*, 2795–2798.

Gyaneshwar, P., Hirsch, A.M., Moulin, L., Chen, W.-M., Elliott, G.N., Bontemps, C., Estrada-de Los Santos, P., Gross, E., Dos Reis, F.B., Sprent, J.I., et al. (2011). Legume-nodulating betaproteobacteria: diversity, host range, and future prospects. Mol. Plant. Microbe. Interact. *24*, 1276–1288.

Haapalainen, M., Mosorin, H., Dorati, F., Wu, R.-F., Roine, E., Taira, S., Nissinen, R., Mattinen, L., Jackson, R., Pirhonen, M., et al. (2012). Hcp2, a secreted protein of the phytopathogen *Pseudomonas syringae* pv. *tomato* DC3000, is required for fitness for competition against bacteria and yeasts. J. Bacteriol. *194*, 4810–4822.

Hachani, A., Lossi, N.S., Hamilton, A., Jones, C., Bleves, S., Albesa-Jové, D., and Filloux, A. (2011). Type VI secretion system in *Pseudomonas aeruginosa*: secretion and multimerization of VgrG proteins. J. Biol. Chem. *286*, 12317–12327.

Haft, D.H., Selengut, J., Mongodin, E.F., and Nelson, K.E. (2005). A guild of 45 CRISPR-associated (Cas) protein families and multiple CRISPR/Cas subtypes exist in prokaryotic genomes. PLoS Comput. Biol. *1*, e60.

Ham, J.H., Majerczak, D.R., Nomura, K., Mecey, C., Uribe, F., He, S.-Y., Mackey, D., and Coplin, D.L. (2009). Multiple activities of the plant pathogen type III effector proteins WtsE and AvrE require WxxxE motifs. Mol. Plant. Microbe. Interact. *22*, 703–712.

Hassan, M., Kjos, M., Nes, I.F., Diep, D.B., and Lotfipour, F. (2012). Natural antimicrobial peptides from bacteria: characteristics and potential applications to fight against antibiotic resistance. J. Appl. Microbiol. *113*, 723–736.

Hess, W.R. (2004). Genome analysis of marine photosynthetic microbes and their global role. Curr. Opin. Biotechnol. *15*, 191–198.

Hockett, K.L., Burch, A.Y., and Lindow, S.E. (2013). Thermo-regulation of genes mediating motility and plant interactions in *Pseudomonas syringae*. PLoS One *8*, e59850.

Höfte, M., and De Vos, P. (2006). Plant pathogenic *Pseudomonas* species. In Plant-Associated Bacteria, S.S. Gnanamanickam, ed. (Dordrecht: Springer Netherlands), pp. 507–533.

Hood, R.D., Singh, P., Hsu, F., Güvener, T., Carl, M.A., Trinidad, R.R.S., Silverman, J.M., Ohlson, B.B., Hicks, K.G., Plemel, R.L., et al. (2010). A type VI secretion system of *Pseudomonas aeruginosa* targets a toxin to bacteria. Cell Host Microbe *7*, 25–37.

Hopwood, D.A. (1997). Genetic Contributions to Understanding Polyketide Synthases. Chem. Rev. *97*, 2465–2498.

Hwang, M.S.H., Morgan, R.L., Sarkar, S.F., Wang, P.W., and Guttman, D.S. (2005). Phylogenetic characterization of virulence and resistance phenotypes of *Pseudomonas syringae*. Appl. Environ. Microbiol. *71*, 5182–5191.

Hyatt, D., Chen, G.-L., Locascio, P.F., Land, M.L., Larimer, F.W., and Hauser, L.J. (2010). Prodigal: prokaryotic gene recognition and translation initiation site identification. BMC Bioinformatics *11*, 119.

Imker, H.J., Walsh, C.T., and Wuest, W.M. (2009). SylC catalyzes ureido-bond formation during biosynthesis of the proteasome inhibitor syringolin A. J. Am. Chem. Soc. *131*, 18263–18265.

Imker, H.J., Krahn, D., Clerc, J., Kaiser, M., and Walsh, C.T. (2010). N-acylation during glidobactin biosynthesis by the tridomain nonribosomal peptide synthetase module GlbF. Chem. Biol. *17*, 1077–1083.

Ishiyama, Y., Yamagishi, N., Ogiso, H., Fujinaga, M., and Takikawa, Y. (2013). Bacterial brown spot on *Avena storigosa* Schereb. caused by *Pseudomonas syringae* pv. *alisalensis*. J. Gen. Plant Pathol. *79*, 155–157.

Jackson, K.D., Starkey, M., Kremer, S., Parsek, M.R., and Wozniak, D.J. (2004). Identification of *psl*, a locus encoding a potential exopolysaccharide that is essential for *Pseudomonas aeruginosa* PAO1 biofilm formation. J. Bacteriol. *186*, 4466–4475.

Jacob, F., Vernaldi, S., and Maekawa, T. (2013). Evolution and Conservation of Plant NLR Functions. Front. Immunol. *4*.

Jelenska, J., Yao, N., Vinatzer, B.A., Wright, C.M., Brodsky, J.L., and Greenberg, J.T. (2007). A J domain virulence effector of *Pseudomonas syringae* remodels host chloroplasts and suppresses defenses. Curr. Biol. *17*, 499–508.

Jelenska, J., van Hal, J.A., and Greenberg, J.T. (2010). *Pseudomonas syringae* hijacks plant stress chaperone machinery for virulence. Proc. Natl. Acad. Sci. U. S. A. *107*, 13177–13182.

Ji, C., Okinaka, Y., Takeuchi, Y., Tsurushima, T., Buzzell, R.I., Sims, J.J., Midland, S.L., Slaymaker, D., Yoshikawa, M., Yamaoka, N., et al. (1997). Specific Binding of the Syringolide Elicitors to a Soluble Protein Fraction from Soybean Leaves. Plant Cell *9*, 1425–1433.

Jiang, H., Rajski, S.R., and Shen, B. (2009). Tandem acyl carrier protein domains in polyunsaturated fatty acid synthases. Methods Enzymol. *459*, 79–96.

Jiang, S., Yao, J., Ma, K.-W., Zhou, H., Song, J., He, S.Y., and Ma, W. (2013). Bacterial Effector Activates Jasmonate Signaling by Directly Targeting JAZ Transcriptional Repressors. PLoS Pathog. *9*, e1003715.

Jiao, N., and Zheng, Q. (2011). The microbial carbon pump: from genes to ecosystems. Appl. Environ. Microbiol. *77*, 7439–7444.

Joardar, V., Lindeberg, M., Jackson, R.W., Selengut, J., Dodson, R., Brinkac, L.M., Daugherty, S.C., Deboy, R., Durkin, A.S., Giglio, M.G., et al. (2005). Whole-genome sequence analysis of *Pseudomonas syringae* pv. *phaseolicola* 1448A reveals divergence among pathovars in genes involved in virulence and transposition. J. Bacteriol. *187*, 6488–6498.

Jones, J.D.G., and Dangl, J.L. (2006). The plant immune system. Nature *444*, 323–329.

Jones, R.K., and Benson, D.M. (2001). Diseases of woody ornamentals and trees in nurseries (St. Paul, MN, USA: APS Press).

De Jong, A., van Heel, A.J., Kok, J., and Kuipers, O.P. (2010). BAGEL2: mining for bacteriocins in genomic data. Nucleic Acids Res. *38*, W647–51.

Joyce, S.A., Watson, R.J., and Clarke, D.J. (2006). The regulation of pathogenicity and mutualism in *Photorhabdus*. Curr. Opin. Microbiol. *9*, 127–132.

Juhas, M., Crook, D.W., and Hood, D.W. (2008). Type IV secretion systems: tools of bacterial horizontal gene transfer and virulence. Cell. Microbiol. *10*, 2377–2386.

Juhas, M., van der Meer, J.R., Gaillard, M., Harding, R.M., Hood, D.W., and Crook, D.W. (2009). Genomic islands: tools of bacterial horizontal gene transfer and evolution. FEMS Microbiol. Rev. *33*, 376–393.

Kämpfer, P. (1997). Detection and cultivation of filamentous bacteria from activated sludge. FEMS Microbiol. Ecol. *23*, 169–181.

Kapitein, N., and Mogk, A. (2013). Deadly syringes: type VI secretion system activities in pathogenicity and interbacterial competition. Curr. Opin. Microbiol. *16*, 52–58.

Keane, P., Kerr, A., and New, P. Crown Gall of Stone Fruit II. Identification and Nomenclature of *Agrobacterium* Isolates. Aust. J. Biol. Sci. *23*, 585–596.

Keen, N.T., Tamaki, S., Kobayashi, D., Gerhold, D., Stayton, M., Shen, H., Gold, S., Lorang, J., Thordal-Christensen, H., Dahlbeck, D., et al. (1990). Bacteria Expressing Avirulence Gene D Produce a Specific Elicitor of the Soybean Hypersensitive Reaction. Mol. Plant-Microbe Interact. *3*, 112–121.

Kelley, D.R., and Estelle, M. (2012). Ubiquitin-mediated control of plant hormone signaling. Plant Physiol. *160*, 47–55.

Kessler, B., de Lorenzo, V., and Timmis, K.N. (1992). A general system to integrate *lacZ* fusions into the chromosomes of gram-negative eubacteria: regulation of the Pm promoter of the TOL plasmid studied with all controlling elements in monocopy. Mol. Gen. Genet. *233*, 293–301.

King, E., Ward, M., and Raney, D. (1954). Two simple media for the demonstration of pyocyanin and fluorescin. J. Lab. Clin. Med. *44*, 301–307.

Kitten, T., Kinscherf, T.G., McEvoy, J.L., and Willis, D.K. (1998). A newly identified regulator is required for virulence and toxin production in *Pseudomonas syringae*. Mol. Microbiol. *28*, 917–929.

Krahn, D., Ottmann, C., and Kaiser, M. (2011). The chemistry and biology of syringolins, glidobactins and cepafungins (syrbactins). Nat. Prod. Rep. *28*, 1854–1867.

Lapouge, K., Schubert, M., Allain, F.H.-T., and Haas, D. (2008). Gac/Rsm signal transduction pathway of gamma-proteobacteria: from RNA recognition to regulation of social behaviour. Mol. Microbiol. *67*, 241–253.

Laue, H., Schenk, A., Li, H., Lambertsen, L., Neu, T.R., Molin, S., and Ullrich, M.S. (2006). Contribution of alginate and levan production to biofilm formation by *Pseudomonas syringae*. Microbiology *152*, 2909–2918.

Lautru, S., Oves-Costales, D., Pernodet, J.-L., and Challis, G.L. (2007). MbtH-like protein-mediated cross-talk between non-ribosomal peptide antibiotic and siderophore biosynthetic pathways in *Streptomyces coelicolor* M145. Microbiology *153*, 1405–1412.

Lee, A.H.-Y., Hurley, B., Felsensteiner, C., Yea, C., Ckurshumova, W., Bartetzko, V., Wang, P.W., Quach, V., Lewis, J.D., Liu, Y.C., et al. (2012a). A bacterial acetyltransferase destroys plant microtubule networks and blocks secretion. PLoS Pathog. *8*, e1002523.

Lee, A.H.-Y., Middleton, M.A., Guttman, D.S., and Desveaux, D. (2013). Phytopathogen type III effectors as probes of biological systems. Microb. Biotechnol. *6*, 230–240.

Lee, J., Teitzel, G.M., Munkvold, K., del Pozo, O., Martin, G.B., Michelmore, R.W., and Greenberg, J.T. (2012b). Type III secretion and effectors shape the survival and growth pattern of *Pseudomonas syringae* on leaf surfaces. Plant Physiol. *158*, 1803–1818.

Lee, S., Choi, H., Suh, S., Doo, I.S., Oh, K.Y., Choi, E.J., Schroeder Taylor, A.T., Low, P.S., and Lee, Y. (1999). Oligogalacturonic acid and chitosan reduce stomatal aperture by inducing the evolution of reactive oxygen species from guard cells of tomato and *Commelina communis*. Plant Physiol. *121*, 147–152.

Van der Lelie, D., Taghavi, S., Monchy, S., Schwender, J., Miller, L., Ferrieri, R., Rogers, A., Wu, X., Zhu, W., Weyens, N., et al. (2009). Poplar and its Bacterial Endophytes: Coexistence and Harmony. CRC. Crit. Rev. Plant Sci. *28*, 346–358.

Lertpiriyapong, K., Gamazon, E.R., Feng, Y., Park, D.S., Pang, J., Botka, G., Graffam, M.E., Ge, Z., and Fox, J.G. (2012). Campylobacter jejuni type VI secretion system: roles in adaptation to deoxycholic acid, host cell adherence, invasion, and in vivo colonization. PLoS One *7*, e42842.

Li, H., and Ullrich, M.S. (2001). Characterization and mutational analysis of three allelic *lsc* genes encoding levansucrase in *Pseudomonas syringae*. J. Bacteriol. *183*, 3282–3292.

Lindeberg, M., Myers, C.R., Collmer, A., and Schneider, D.J. (2008). Roadmap to new virulence determinants in *Pseudomonas syringae*: insights from comparative genomics and genome organization. Mol. Plant. Microbe. Interact. *21*, 685–700.

Lindeberg, M., Cunnac, S., and Collmer, A. (2009). The evolution of *Pseudomonas syringae* host specificity and type III effector repertoires. Mol. Plant Pathol. *10*, 767–775.

Lindeberg, M., Cunnac, S., and Collmer, A. (2012). *Pseudomonas syringae* type III effector repertoires: last words in endless arguments. Trends Microbiol. *20*, 199–208.

Liu, H., Qiu, H., Zhao, W., Cui, Z., Ibrahim, M., Jin, G., Li, B., Zhu, B., and Xie, G.L. (2012). Genome sequence of the plant pathogen *Pseudomonas syringae* pv. *panici* LMG 2367. J. Bacteriol. *194*, 5693–5694.

Long, S.R. (1996). Rhizobium symbiosis: nod factors in perspective. Plant Cell *8*, 1885–1898.

Lu, S.-E., Scholz-Schroeder, B.K., and Gross, D.C. (2002). Characterization of the salA, syrF, and syrG Regulatory Genes Located at the Right Border of the Syringomycin Gene Cluster of Pseudomonas syringae pv. syringae. Mol. Plant-Microbe Interact. *15*, 43–53.

Mabood, F., and Smith, D. (2007). The Role of Salicylates in Rhizobium-Legume Symbiosis and Abiotic Stresses in Higher Plants. In Salicylic Acid: A Plant Hormone, S. Hayat, and A. Ahmad, eds. (Dordrecht, The Netherlands: Springer Netherlands), pp. 151–162.

Mackey, D., Belkhadir, Y., Alonso, J.M., Ecker, J.R., and Dangl, J.L. (2003). *Arabidopsis* RIN4 Is a Target of the Type III Virulence Effector AvrRpt2 and Modulates RPS2-Mediated Resistance. Cell *112*, 379–389.

Makarova, K.S., Wolf, Y.I., van der Oost, J., and Koonin, E. V (2009). Prokaryotic homologs of Argonaute proteins are predicted to function as key components of a novel system of defense against mobile genetic elements. Biol. Direct *4*, 29.

Makarova, K.S., Wolf, Y.I., and Koonin, E. V (2013). Comparative genomics of defense systems in archaea and bacteria. Nucleic Acids Res. *41*, 4360–4377.

Mansfield, J., Genin, S., Magori, S., Citovsky, V., Sriariyanum, M., Ronald, P., Dow, M., Verdier, V., Beer, S. V, Machado, M.A., et al. (2012). Top 10 plant pathogenic bacteria in molecular plant pathology. Mol. Plant Pathol. *13*, 614–629.

Marcelletti, S., Ferrante, P., Petriccione, M., Firrao, G., and Scortichini, M. (2011). *Pseudomonas syringae* pv. *actinidiae* draft genomes comparison reveal strain-specific features involved in adaptation and virulence to *Actinidia* species. PLoS One *6*, e27297.

Marchler-Bauer, A., Lu, S., Anderson, J.B., Chitsaz, F., Derbyshire, M.K., DeWeese-Scott, C., Fong, J.H., Geer, L.Y., Geer, R.C., Gonzales, N.R., et al. (2011). CDD: a Conserved Domain Database for the functional annotation of proteins. Nucleic Acids Res. *39*, D225–9.

Markowitz, V.M., Mavromatis, K., Ivanova, N.N., Chen, I.-M.A., Chu, K., and Kyrpides, N.C. (2009). IMG ER: a system for microbial genome annotation expert review and curation. Bioinformatics *25*, 2271–2278.

Martín-Sanz, A., de la Vega, M.P., Murillo, J., and Caminero, C. (2013). Strains of *Pseudomonas syringae* pv. *syringae* from pea are phylogenetically and pathogenically diverse. Phytopathology *103*, 673–681.

Matas, I.M., Castañeda-Ojeda, M.P., Aragón, I.M., Antúnez-Lamas, M., Murillo, J., Rodriquez-Palenzuela, P., Lopez-Solanilla, E., and Ramos, C. (2014). Translocation and functional analysis of Pseudomonas savastanoi pv. savastanoi NCPPB 3335 type III secretion system effectors reveals two novel effector families of the Pseudomonas syringae complex. Mol. Plant. Microbe. Interact. In press.

Matsukawa, M., and Greenberg, E.P. (2004). Putative exopolysaccharide synthesis genes influence *Pseudomonas aeruginosa* biofilm development. J. Bacteriol. *186*, 4449–4456.

McCann, H.C., Rikkerink, E.H.A., Bertels, F., Fiers, M., Lu, A., Rees-George, J., Andersen, M.T., Gleave, A.P., Haubold, B., Wohlers, M.W., et al. (2013). Genomic Analysis of the Kiwifruit Pathogen *Pseudomonas syringae* pv. *actinidiae* Provides Insight into the Origins of an Emergent Plant Disease. PLoS Pathog. *9*, e1003503.

Meier-Kolthoff, J.P., Auch, A.F., Klenk, H.-P., and Göker, M. (2013). Genome sequence-based species delimitation with confidence intervals and improved distance functions. BMC Bioinformatics *14*, 60.

Melotto, M., Underwood, W., Koczan, J., Nomura, K., and He, S.Y. (2006). Plant Stomata Function in Innate Immunity against Bacterial Invasion. Cell *126*, 969–980.

Melotto, M., Underwood, W., and He, S.Y. (2008). Role of stomata in plant innate immunity and foliar bacterial diseases. Annu. Rev. Phytopathol. *46*, 101–122.

Miché, L., Battistoni, F., Gemmer, S., Belghazi, M., and Reinhold-Hurek, B. (2006). Upregulation of jasmonate-inducible defense proteins and differential colonization of roots of *Oryza sativa* cultivars with the endophyte *Azoarcus sp*. Mol. Plant. Microbe. Interact. *19*, 502–511.

Miller, M.B., and Bassler, B.L. (2001). Quorum sensing in bacteria. Annu. Rev. Microbiol. *55*, 165–199.

Mima, T., Sekiya, H., Mizushima, T., Kuroda, T., and Tsuchiya, T. (2005). Gene cloning and properties of the RND-type multidrug efflux pumps MexPQ-OpmE and MexMN-OprM from *Pseudomonas aeruginosa*. Microbiol. Immunol. *49*, 999–1002.

Misas-Villamil, J.C., Kolodziejek, I., Crabill, E., Kaschani, F., Niessen, S., Shindo, T., Kaiser, M., Alfano, J.R., and van der Hoorn, R.A.L. (2013). *Pseudomonas syringae* pv. *syringae* uses proteasome inhibitor syringolin A to colonize from wound infection sites. PLoS Pathog. *9*, e1003281.

Mo, Y.Y., and Gross, D.C. (1991). Plant signal molecules activate the *syrB* gene, which is required for syringomycin production by *Pseudomonas syringae* pv. *syringae*. J. Bacteriol. *173*, 5784–5792.

Mucyn, T.S., Yourstone, S., Lind, A.L., Biswas, S., Nishimura, M.T., Baltrus, D.A., Cumbie, J.S., Chang, J.H., Jones, C.D., Dangl, J.L., et al. (2014). Variable Suites of Non-effector Genes Are Co-regulated in the Type III Secretion Virulence Regulon across the *Pseudomonas syringae* Phylogeny. PLoS Pathog. *10*, e1003807.

Mulder, D.T., Cooper, C.A., and Coombes, B.K. (2012). Type VI secretion system-associated gene clusters contribute to pathogenesis of *Salmonella enterica* serovar Typhimurium. Infect. Immun. *80*, 1996–2007.

Munkvold, K.R., Martin, M.E., Bronstein, P.A., and Collmer, A. (2008). A survey of the *Pseudomonas syringae* pv. *tomato* DC3000 type III secretion system effector repertoire reveals several effectors that are deleterious when expressed in *Saccharomyces cerevisiae*. Mol. Plant-Microbe Interact. *21*, 490–502.

Munkvold, K.R., Russell, A.B., Kvitko, B.H., and Collmer, A. (2009). *Pseudomonas syringae* pv. *tomato* DC3000 type III effector HopAA1-1 functions redundantly with chlorosis-promoting factor PSPTO4723 to produce bacterial speck lesions in host tomato. Mol. Plant. Microbe. Interact. *22*, 1341–1355.

Murillo, J., Bardaji, L., Navarro de la Fuente, L., Elena Führer, M., Aguilera, S., and Álvarez-Morales, A. (2011). Variation in conservation of the cluster for biosynthesis of the phytotoxin phaseolotoxin in *Pseudomonas syringae* suggests at least two events of horizontal acquisition. Res. Microbiol. *162*, 253–261.

Mysore, K.S., and Ryu, C.-M. (2004). Nonhost resistance: how much do we know? Trends Plant Sci. *9*, 97–104.

Nealson, K.H. (1977). Autoinduction of bacterial luciferase. Arch. Microbiol. *112*, 73–79.

Newman, D.J., Cragg, G.M., and Snader, K.M. (2003). Natural products as sources of new drugs over the period 1981-2002. J. Nat. Prod. *66*, 1022–1037.

Nicaise, V., Joe, A., Jeong, B.-R., Korneli, C., Boutrot, F., Westedt, I., Staiger, D., Alfano, J.R., and Zipfel, C. (2013). *Pseudomonas* HopU1 modulates plant immune receptor levels by blocking the interaction of their mRNAs with GRP7. EMBO J. *32*, 701–712.

Nikolouli, K., and Mossialos, D. (2012). Bioactive compounds synthesized by non-ribosomal peptide synthetases and type-I polyketide synthases discovered through genome-mining and metagenomics. Biotechnol. Lett. *34*, 1393–1403.

Nomura, K., Debroy, S., Lee, Y.H., Pumplin, N., Jones, J., and He, S.Y. (2006). A bacterial virulence protein suppresses host innate immunity to cause plant disease. Science *313*, 220–223.

Nomura, K., Mecey, C., Lee, Y.-N., Imboden, L.A., Chang, J.H., and He, S.Y. (2011). Effector-triggered immunity blocks pathogen degradation of an immunity-associated vesicle traffic regulator in *Arabidopsis*. Proc. Natl. Acad. Sci. U. S. A. *108*, 10774–10779.

O'Brien, H.E., Thakur, S., and Guttman, D.S. (2011). Evolution of plant pathogenesis in *Pseudomonas syringae*: a genomics perspective. Annu. Rev. Phytopathol. *49*, 269–289.

O'Brien, H.E., Thakur, S., Gong, Y., Fung, P., Zhang, J., Yuan, L., Wang, P.W., Yong, C., Scortichini, M., and Guttman, D.S. (2012). Extensive remodeling of the *Pseudomonas syringae* pv. *avellanae* type III secretome associated with two independent host shifts onto hazelnut. BMC Microbiol. *12*, 141.

Oerke, E.-C., and Dehne, H.-W. (2004). Safeguarding production—losses in major crops and the role of crop protection. Crop Prot. *23*, 275–285.

Oka, M., Nishiyama, Y., Ohta, S., Kamei, H., Konishi, M., Miyaki, T., Oki, T., and Kawaguchi, H. (1988a). Glidobactins A, B and C, new antitumor antibiotics. I. Production, isolation, chemical properties and biological activity. J. Antibiot. (Tokyo). *41*, 1331–1337.

Oka, M., Yaginuma, K., Numata, K., Konishi, M., Oki, T., and Kawaguchi, H. (1988b). Glidobactins A, B and C, new antitumor antibiotics. II. Structure elucidation. J. Antibiot. (Tokyo). *41*, 1338–1350.

Oka, M., Ohkuma, H., Kamei, H., Konishi, M., Oki, T., and Kawaguchi, H. (1988c). Glidobactins D, E, F, G and H; minor components of the antitumor antibiotic glidobactin. J. Antibiot. (Tokyo). *41*, 1906–1909.

Orcutt, B.N., Sylvan, J.B., Knab, N.J., and Edwards, K.J. (2011). Microbial ecology of the dark ocean above, at, and below the seafloor. Microbiol. Mol. Biol. Rev. *75*, 361–422.

Osman, S.F., Fett, W.F., and Fishman, M.L. (1986). Exopolysaccharides of the phytopathogen *Pseudomonas syringae* pv. *glycinea*. J. Bacteriol. *166*, 66–71.

Owen, J.G., and Ackerley, D.F. (2011). Characterization of pyoverdine and achromobactin in *Pseudomonas syringae* pv. *phaseolicola* 1448a. BMC Microbiol. *11*.

Palleroni, N. (2005a). Genus I. Pseudomonas. In Bergey's Manual of Systematic Bacteriology, Second Edition, Volume 2, Part B, G. Garrity, D. Brenner, N. Krieg, and J. Staley, eds. (Springer), p. 374.

Palleroni, N. (2005b). Genus I. Pseudomonas. In Bergey's Manual of Systematic Bacteriology, Second Edition, Volume 2, Part B, G. Garrity, D. Brenner, N. Krieg, and J. Staley, eds. (Springer, New York), pp. 323–379.

Palmer, D.A., and Bender, C.L. (1993). Effects of Environmental and Nutritional Factors on Production of the Polyketide Phytotoxin Coronatine by *Pseudomonas syringae* pv. *Glycinea*. Appl. Envir. Microbiol. *59*, 1619–1626.

Penaloza-Vazquez, A., Kidambi, S., Chakrabarty, A., and Bender, C. (1997). Characterization of the alginate biosynthetic gene cluster in *Pseudomonas syringae* pv. *syringae*. J. Bacteriol. *179*, 4464–4472.

Pérez-Martínez, I., Zhao, Y., Murillo, J., Sundin, G.W., and Ramos, C. (2008). Global genomic analysis of *Pseudomonas savastanoi* pv. *savastanoi* plasmids. J. Bacteriol. *190*, 625–635.

Philippot, L., Raaijmakers, J.M., Lemanceau, P., and van der Putten, W.H. (2013). Going back to the roots: the microbial ecology of the rhizosphere. Nat. Rev. Microbiol. *11*, 789–799.

Pieterse, C.M.J., and Dicke, M. (2007). Plant interactions with microbes and insects: from molecular mechanisms to ecology. Trends Plant Sci. *12*, 564–569.

Pitman, A.R., Jackson, R.W., Mansfield, J.W., Kaitell, V., Thwaites, R., and Arnold, D.L. (2005). Exposure to host resistance mechanisms drives evolution of bacterial virulence in plants. Curr. Biol. *15*, 2230–2235.

Poelarends, G.J., Veetil, V.P., and Whitman, C.P. (2008). The chemical versatility of the beta-alpha-beta fold: catalytic promiscuity and divergent evolution in the tautomerase superfamily. Cell. Mol. Life Sci. *65*, 3606–3618.

Poole, K. (2001). Multidrug efflux pumps and antimicrobial resistance in *Pseudomonas aeruginosa* and related organisms. J. Mol. Microbiol. Biotechnol. *3*, 255–264.

Prieto, C., García-Estrada, C., Lorenzana, D., and Martín, J.F. (2012). NRPSsp: non-ribosomal peptide synthase substrate predictor. Bioinformatics *28*, 426–427.

Pukatzki, S., Ma, A.T., Revel, A.T., Sturtevant, D., and Mekalanos, J.J. (2007). Type VI secretion system translocates a phage tail spike-like protein into target cells where it cross-links actin. Proc. Natl. Acad. Sci. U. S. A. *104*, 15508–15513.

Qi, D., and Innes, R.W. (2013). Recent Advances in Plant NLR Structure, Function, Localization, and Signaling. Front. Immunol. *4*, 348.

Qi, M., Wang, D., Bradley, C.A., and Zhao, Y. (2011). Genome sequence analyses of *Pseudomonas savastanoi* pv. *glycinea* and subtractive hybridization-based comparative genomics with nine Pseudomonads. PLoS One *6*, e16451.

Quandt, J., and Hynes, M.F. (1993). Versatile suicide vectors which allow direct selection for gene replacement in gram-negative bacteria. Gene *127*, 15–21.

Rackham, E.J., Grüschow, S., Ragab, A.E., Dickens, S., and Goss, R.J.M. (2010). Pacidamycin biosynthesis: identification and heterologous expression of the first uridyl peptide antibiotic gene cluster. Chembiochem *11*, 1700–1709.

Rahimian, H. (1995). The Occurrence of Bacterial Red Streak of Sugarcane Caused by *Pseudomonas syringae* pv. *syringae* in Iran. J. Phytopathol. *143*, 321–324.

Ramel, C., Tobler, M., Meyer, M., Bigler, L., Ebert, M.-O., Schellenberg, B., and Dudler, R. (2009). Biosynthesis of the proteasome inhibitor syringolin A: the ureido group joining two amino acids originates from bicarbonate. BMC Biochem. *10*, 26.

Ramel, C., Baechler, N., Hildbrand, M., Meyer, M., Schädeli, D., and Dudler, R. (2012). Regulation of biosynthesis of syringolin A, a *Pseudomonas syringae* virulence factor targeting the host proteasome. Mol. Plant-Microbe Interact. *25*, 1198–1208.

Ramos, C., Matas, I.M., Bardaji, L., Aragón, I.M., and Murillo, J. (2012). *Pseudomonas savastanoi* pv. *savastanoi*: some like it knot. Mol. Plant Pathol. *13*, 998–1009.

Rao, P.S.S., Yamada, Y., Tan, Y.P., and Leung, K.Y. (2004). Use of proteomics to identify novel virulence determinants that are required for *Edwardsiella tarda* pathogenesis. Mol. Microbiol. *53*, 573–586.

Rasmussen, J.B., Hammerschmidt, R., and Zook, M.N. (1991). Systemic Induction of Salicylic Acid Accumulation in Cucumber after Inoculation with *Pseudomonas syringae* pv. *syringae*. Plant Physiol. *97*, 1342–1347.

Records, A.R., and Gross, D.C. (2010). Sensor kinases RetS and LadS regulate *Pseudomonas syringae* type VI secretion and virulence factors. J. Bacteriol. *192*, 3584–3596.

Reimmann, C., Hofmann, C., Mauch, F., and Dudler, R. (1995). Characterization of a rice gene induced by *Pseudomonas syringae* pv. *syringae*: requirement for the bacterial *lemA* gene function. Physiol. Mol. Plant Pathol. *46*, 71–81.

Reinhardt, J.A., Baltrus, D.A., Nishimura, M.T., Jeck, W.R., Jones, C.D., and Dangl, J.L. (2009). De novo assembly using low-coverage short read sequence data from the rice pathogen *Pseudomonas syringae* pv. *oryzae*. Genome Res. *19*, 294–305.

Reinhold-Hurek, B., and Hurek, T. (2011). Living inside plants: bacterial endophytes. Curr. Opin. Plant Biol. *14*, 435–443.

Roberto, F.F., Klee, H., White, F., Nordeen, R., and Kosuge, T. (1990). Expression and fine structure of the gene encoding N epsilon-(indole-3-acetyl)-L-lysine synthetase from *Pseudomonas savastanoi*. Proc. Natl. Acad. Sci. U. S. A. *87*, 5797–5801.

Rodríguez-Blanco, A., Lemos, M.L., and Osorio, C.R. (2012). Integrating conjugative elements as vectors of antibiotic, mercury, and quaternary ammonium compound resistance in marine aquaculture environments. Antimicrob. Agents Chemother. *56*, 2619–2626.

Rodríguez-Herva, J.J., González-Melendi, P., Cuartas-Lanza, R., Antúnez-Lamas, M., Río-Alvarez, I., Li, Z., López-Torrejón, G., Díaz, I., Del Pozo, J.C., Chakravarthy, S., et al. (2012). A bacterial cysteine protease effector protein interferes with photosynthesis to suppress plant innate immune responses. Cell. Microbiol. *14*, 669–681.

Rodríguez-Moreno, L., Barceló-Muñoz, A., and Ramos, C. (2008). In vitro analysis of the interaction of *Pseudomonas savastanoi* pvs. *savastanoi* and *nerii* with micropropagated olive plants. Phytopathology *98*, 815–822.

Rodríguez-Palenzuela, P., Matas, I.M., Murillo, J., López-Solanilla, E., Bardaji, L., Pérez-Martínez, I., Rodríguez-Moskera, M.E., Penyalver, R., López, M.M., Quesada, J.M., et al. (2010). Annotation and overview of the *Pseudomonas savastanoi* pv. *savastanoi* NCPPB 3335 draft genome reveals the

virulence gene complement of a tumour-inducing pathogen of woody hosts. Environ. Microbiol. *12*, 1604–1620.

Rosebrock, T.R., Zeng, L., Brady, J.J., Abramovitch, R.B., Xiao, F., and Martin, G.B. (2007). A bacterial E3 ubiquitin ligase targets a host protein kinase to disrupt plant immunity. Nature *448*, 370–374.

Rosenblueth, M., and Martínez-Romero, E. (2006). Bacterial endophytes and their interactions with hosts. Mol. Plant. Microbe. Interact. *19*, 827–837.

Russell, A.B., LeRoux, M., Hathazi, K., Agnello, D.M., Ishikawa, T., Wiggins, P.A., Wai, S.N., and Mougous, J.D. (2013). Diverse type VI secretion phospholipases are functionally plastic antibacterial effectors. Nature *496*, 508–512.

Salomon, D., Bosis, E., Dar, D., Nachman, I., and Sessa, G. (2012). Expression of *Pseudomonas syringae* type III effectors in yeast under stress conditions reveals that HopX1 attenuates activation of the high osmolarity glycerol MAP kinase pathway. Microbiology *158*, 2859–2869.

Sambrook, J., and Russell, D.W. (2001). Molecular Cloning: A Laboratory Manual. (Cold Spring Harbor, NY, USA: Cold Spring Harbor Laboratory Press).

Sarkar, S.F., and Guttman, D.S. (2004). Evolution of the core genome of *Pseudomonas syringae*, a highly clonal, endemic plant pathogen. Appl. Environ. Microbiol. *70*, 1999–2012.

Sarkar, S.F., Gordon, J.S., Martin, G.B., and Guttman, D.S. (2006). Comparative genomics of host-specific virulence in *Pseudomonas syringae*. Genetics *174*, 1041–1056.

Sarris, P.F., Skandalis, N., Kokkinidis, M., and Panopoulos, N.J. (2010). *In silico* analysis reveals multiple putative type VI secretion systems and effector proteins in *Pseudomonas syringae* pathovars. Mol. Plant Pathol. *11*, 795–804.

Sarris, P.F., Trantas, E.A., Skandalis, N., Tampakaki, A.P., Kapanidou, M., Kokkinidis, M., and Panopoulos, N.J. (2012). Phytobacterial Type VI Secretion System - Gene Distribution, Phylogeny, Structure and Biological Functions. In InTech Plant Pathology, C.J.R. Cumagun, ed. (Rijeka, Croatia: InTech), pp. 53–84.

Sarris, P.F., Trantas, E.A., Baltrus, D.A., Bull, C.T., Wechter, W.P., Yan, S., Ververidis, F., Almeida, N.F., Jones, C.D., Dangl, J.L., et al. (2013). Comparative genomics of multiple strains of *Pseudomonas cannabina* pv. *alisalensis*, a potential model pathogen of both monocots and dicots. PLoS One *8*, e59366.

Saxena, A.K., Pal, K.K., and Tilak, K.V.B.R. (2000). Bacterial biocontrol agents and their role in plant disease management. In Biocontrol Potential and Its Exploitation in Sustainable Agriculture.

Volume 1: Crop Diseases, Weeds, and Nematodes, R.K. Upadhyay, K.G. Mukerji, and B.P. Chamola, eds. (New York, USA: Springer US), pp. 25–37.

Schellenberg, B., Bigler, L., and Dudler, R. (2007). Identification of genes involved in the biosynthesis of the cytotoxic compound glidobactin from a soil bacterium. Environ. Microbiol. *9*, 1640–1650.

Schellenberg, B., Ramel, C., and Dudler, R. (2010). *Pseudomonas syringae* Virulence Factor Syringolin A Counteracts Stomatal Immunity by Proteasome Inhibition. Mol. Plant-Microbe Interact. *23*, 1287–1293.

Schikora, A., Schenk, S.T., Stein, E., Molitor, A., Zuccaro, A., and Kogel, K.-H. (2011). N-acyl-homoserine lactone confers resistance toward biotrophic and hemibiotrophic pathogens via altered activation of AtMPK6. Plant Physiol. *157*, 1407–1418.

Schnider-Keel, U., Lejbølle, K.B., Baehler, E., Haas, D., and Keel, C. (2001). The sigma factor AlgU (AlgT) controls exopolysaccharide production and tolerance towards desiccation and osmotic stress in the biocontrol agent *Pseudomonas fluorescens* CHA0. Appl. Environ. Microbiol. *67*, 5683–5693.

Schwarzer, D., Mootz, H.D., Linne, U., and Marahiel, M.A. (2002). Regeneration of misprimed nonribosomal peptide synthetases by type II thioesterases. Proc. Natl. Acad. Sci. U. S. A. *99*, 14083–14088.

Seemann, T. (2013). Prokka: Prokaryotic Genome Annotation System.

Senior, E., Bull, A.T., and Slater, J.H. (1976). Enzyme evolution in a microbial community growing on the herbicide Dalapon. Nature *263*, 476–479.

Senthil-Kumar, M., and Mysore, K.S. (2013). Nonhost Resistance Against Bacterial Pathogens: Retrospectives and Prospects. *51*, 407–427.

Sha, J., Rosenzweig, J.A., Kozlova, E. V, Wang, S., Erova, T.E., Kirtley, M.L., van Lier, C.J., and Chopra, A.K. (2013). Evaluation of the roles played by Hcp and VgrG type 6 secretion system effectors in *Aeromonas hydrophila* SSU pathogenesis. Microbiology *159*, 1120–1135.

Shaw, S.L., and Long, S.R. (2003). Nod Factor Inhibition of Reactive Oxygen Efflux in a Host Legume. Plant Physiol. *132*, 2196–2204.

Shrivastava, S., and Mande, S.S. (2008). Identification and functional characterization of gene components of Type VI Secretion system in bacterial genomes. PLoS One *3*, e2955.

Simon, R., Priefer, U., and Pühler, A. (1983). A Broad Host Range Mobilization System for In Vivo Genetic Engineering: Transposon Mutagenesis in Gram Negative Bacteria. Bio/Technology *1*, 784–791.

Sinclair, W.A., Lyon, H.H., and Johnson, W.T. (1987). Diseases of trees and shrubs (Ithaca, NY, USA: Cornell Univrersity Press).

Smith, J.A., and Métraux, J.-P. (1991). *Pseudomonas syringae* pv. *syringae* induces systemic resistance to *Pyricularia oryzae* in rice. Physiol. Mol. Plant Pathol. *39*, 451–461.

Smith, J.A., Hammerschmidt, R., and Fulbright, D.W. (1991). Rapid induction of systemic resistance in cucumber by *Pseudomonas syringae* pv. *syringae*. Physiol. Mol. Plant Pathol. *38*, 223–235.

Stachelhaus, T., and Marahiel, M.A. (1995). Modular structure of genes encoding multifunctional peptide synthetases required for non-ribosomal peptide synthesis. FEMS Microbiol. Lett. *125*, 3–14.

Stachelhaus, T., Mootz, H.D., and Marahiel, M.A. (1999). The specificity-conferring code of adenylation domains in nonribosomal peptide synthetases. Chem. Biol. *6*, 493–505.

Staskawicz, B., Dahlbeck, D., Keen, N., and Napoli, C. (1987). Molecular characterization of cloned avirulence genes from race 0 and race 1 of *Pseudomonas syringae* pv. *glycinea*. J. Bacteriol. *169*, 5789–5794.

Stein, M.L., Beck, P., Kaiser, M., Dudler, R., Becker, C.F.W., and Groll, M. (2012). One-shot NMR analysis of microbial secretions identifies highly potent proteasome inhibitor. Proc. Natl. Acad. Sci. U. S. A. *109*, 18367–18371.

Strieker, M., Tanović, A., and Marahiel, M.A. (2010). Nonribosomal peptide synthetases: structures and dynamics. Curr. Opin. Struct. Biol. *20*, 234–240.

Subramoni, S., and Venturi, V. (2009). LuxR-family "solos": bachelor sensors/regulators of signalling molecules. Microbiology *155*, 1377–1385.

Subramoni, S., Gonzalez, J.F., Johnson, A., Péchy-Tarr, M., Rochat, L., Paulsen, I., Loper, J.E., Keel, C., and Venturi, V. (2011). Bacterial subfamily of LuxR regulators that respond to plant compounds. Appl. Environ. Microbiol. *77*, 4579–4588.

Sundin, G.W., Demezas, D.H., and Bender, C.L. (1994). Genetic and plasmid diversity within natural populations of *Pseudomonas syringae* with various exposures to copper and streptomycin bactericides. Appl. Envir. Microbiol. *60*, 4421–4431.

Swingle, B., Bao, Z., Markel, E., Chambers, A., and Cartinhour, S. (2010). Recombineering using RecTE from *Pseudomonas syringae*. Appl. Environ. Microbiol. *76*, 4960–4968.

Takai, K., and Nakamura, K. (2011). Archaeal diversity and community development in deep-sea hydrothermal vents. Curr. Opin. Microbiol. *14*, 282–291.

Tamura, K., Peterson, D., Peterson, N., Stecher, G., Nei, M., and Kumar, S. (2011). MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. Mol. Biol. Evol. *28*, 2731–2739.

Tang, X., Xiao, Y., and Zhou, J.-M. (2006). Regulation of the type III secretion system in phytopathogenic bacteria. Mol. Plant. Microbe. Interact. *19*, 1159–1166.

Terui, Y., Nishikawa, J., Hinoo, H., Kato, T., and Shoji, J. (1990). Structures of cepafungins I, II and III. J. Antibiot. (Tokyo). *43*, 788–795.

Theodore, C.M., King, J.B., You, J., and Cichewicz, R.H. (2012). Production of cytotoxic glidobactins/luminmycins by *Photorhabdus asymbiotica* in liquid media and live crickets. J. Nat. Prod. *75*, 2007–2011.

Thoma, S., and Schobert, M. (2009). An improved *Escherichia coli* donor strain for diparental mating. FEMS Microbiol. Lett. *294*, 127–132.

Titus, J.A., and Roundy, C.A. (1990). Improvement of glidobactin A production by *Polyangium brachysporum*. J. Ind. Microbiol. *6*, 215–218.

Toussaint, A., Prangishvili, D., Molineux, I., Fineran, P.C., and Charpentier, E. (2012). Memory of viral infections by CRISPR-Cas adaptive immune systems: Acquisition of new information. Virology *434*, 202–209.

Velázquez, E., Palomo, J.L., Rivas, R., Guerra, H., Peix, A., Trujillo, M.E., García-Benavides, P., Mateos, P.F., Wabiko, H., and Martínez-Molina, E. (2010). Analysis of core genes supports the reclassification of strains *Agrobacterium radiobacter* K84 and *Agrobacterium tumefaciens* AKE10 into the species *Rhizobium rhizogenes*. Syst. Appl. Microbiol. *33*, 247–251.

Venturi, V., and Fuqua, C. (2013). Chemical signaling between plants and plant-pathogenic bacteria. Annu. Rev. Phytopathol. *51*, 17–37.

Vincent, J.R., and Fulbright, D.W. (1983). Transfer of pRD1 to *Pseudomonas syringae* and evidence for its integration into the chromosome. J. Bacteriol. *156*, 1349–1351.

Voisard, C., Bull, C.T., Keel, C., Laville, J., Maurhofer, M., Schnider, U., D´efago, G., and Haas, D. (1994). Chapter 6. Biocontrol of Root Diseases by *Pseudomonas fluorescens* CHA0: Current Concepts and Experimental Approaches. In Molecular Ecology of Rhizosphere Microorganisms, F. O'Gara, D.N. Dowling, and B. Boesten, eds. (Weinheim, Germany: Wiley-VCH Verlag GmbH), pp. 67–89.

Vorholt, J.A. (2012). Microbial life in the phyllosphere. Nat. Rev. Microbiol. *10*, 828–840.

Wang, Y., Li, J., Hou, S., Wang, X., Li, Y., Ren, D., Chen, S., Tang, X., and Zhou, J.-M. (2010). A *Pseudomonas syringae* ADP-ribosyltransferase inhibits *Arabidopsis* mitogen-activated protein kinase kinases. Plant Cell *22*, 2033–2044.

Wang, Y., Huang, H., Sun, M., Zhang, Q., and Guo, D. (2012). T3DB: an integrated database for bacterial type III secretion system. BMC Bioinformatics *13*, 66.

Waspi, U. (2001). Syringolin Reprograms Wheat to Undergo Hypersensitive Cell Death in a Compatible Interaction with Powdery Mildew. PLANT CELL ONLINE *13*, 153–161.

Wäspi, U., Misteli, B., Hasslacher, M., Jandrositz, A., Kohlwein, S.D., Schwab, H., and Dudler, R. (1998a). The defense-related rice gene Pir7b encodes an alpha/beta hydrolase fold protein exhibiting esterase activity towards naphthol AS-esters. Eur. J. Biochem. *254*, 32–37.

Wäspi, U., Blanc, D., Winkler, T., Rüedi, P., and Dudler, R. (1998b). Syringolin, a Novel Peptide Elicitor from *Pseudomonas syringae* pv. *syringae* that Induces Resistance to *Pyricularia oryzae* in Rice. Mol. Plant-Microbe Interact. *11*, 727–733.

Wäspi, U., Hassa, P., Staempfli, A.A., Molleyres, L.-P., Winkler, T., and Dudler, R. (1999). Identification and structure of a family of syringolin variants : Unusual cyclic peptides from *Pseudomonas syringae* pv. *syringae* that elicit defense responses in rice. Microbiol. Res. *154*, 89–93.

Wei, C.-F., Kvitko, B.H., Shimizu, R., Crabill, E., Alfano, J.R., Lin, N.-C., Martin, G.B., Huang, H.-C., and Collmer, A. (2007). A *Pseudomonas syringae* pv. *tomato* DC3000 mutant lacking the type III effector HopQ1-1 is able to cause disease in the model plant *Nicotiana benthamiana*. Plant J. *51*, 32–46.

Weingart, H., Völksch, B., and Ullrich, M.S. (1999). Comparison of Ethylene Production by *Pseudomonas syringae* and *Ralstonia solanacearum*. Phytopathology *89*, 360–365.

Whitman, W.B. (1998). Prokaryotes: The unseen majority. Proc. Natl. Acad. Sci. *95*, 6578–6583.

Williamson, C.E. (1979). An Ultrastructural Investigation of Algal Symbiosis in White and Green *Spongilla lacustris* (L.) (Porifera: Spongillidae). Trans. Am. Microsc. Soc. *98*, 59–77.

Willis, D.K., Hrabak, E.M., Rich, J.J., Barta, T.M., Lindow, S.E., and Panopoulos, N.J. (1990). Isolation and Characterization of a *Pseudomonas syringae* pv. *syringae* Mutant Deficient in Lesion Formation on Bean. Mol. Plant-Microbe Interact. *3*, 149.

Woese, C.R. (1990). Towards a Natural System of Organisms: Proposal for the Domains Archaea, Bacteria, and Eucarya. Proc. Natl. Acad. Sci. *87*, 4576–4579.

Wozniak, R.A.F., and Waldor, M.K. (2010). Integrative and conjugative elements: mosaic mobile genetic elements enabling dynamic lateral gene flow. Nat. Rev. Microbiol. *8*, 552–563.

Wuest, W.M., Krahn, D., Kaiser, M., and Walsh, C.T. (2011). Enzymatic timing and tailoring of macrolactamization in syringolin biosynthesis. Org. Lett. *13*, 4518–4521.

Xiang, T., Zong, N., Zou, Y., Wu, Y., Zhang, J., Xing, W., Li, Y., Tang, X., Zhu, L., Chai, J., et al. (2008). *Pseudomonas syringae* Effector AvrPto Blocks Innate Immunity by Targeting Receptor Kinases. Curr. Biol. *18*, 74–80.

Yasuta, T., Okazaki, S., Mitsui, H., Yuhashi, K., Ezura, H., and Minamisawa, K. (2001). DNA sequence and mutational analysis of rhizobitoxine biosynthesis genes in *Bradyrhizobium elkanii*. Appl. Environ. Microbiol. *67*, 4999–5009.

Yeh, E., Kohli, R.M., Bruner, S.D., and Walsh, C.T. (2004). Type II thioesterase restores activity of a NRPS module stalled with an aminoacyl-S-enzyme that cannot be elongated. Chembiochem *5*, 1290–1293.

Young, J.M. (2010). Taxonomy of *Pseudomonas syringae*. J. Plant Pathol. *92*, S1.5–S1.14.

Young, J.M., Luketina, R.C., and Marshall, A.M. (1977). The Effects on Temperature on Growth in vitro of *Pseudomonas syringae* and *Xanthomonas pruni*. J. Appl. Microbiol. *42*, 345–354.

Young, J.M., Dye, D.W., Bradbury, J.F., Panagopoulos, C.G., and Robbs, C.F. (1978). A proposed nomenclature and classification for plant pathogenic bacteria. New Zeal. J. Agric. Res. *21*, 153–177.

Yu, J., Penaloza-Vazquez, A., Chakrabarty, A.M., and Bender, C.L. (1999). Involvement of the exopolysaccharide alginate in the virulence and epiphytic fitness of *Pseudomonas syringae* pv. *syringae*. Mol. Microbiol. *33*, 712–720.

Yu, X., Lund, S.P., Scott, R.A., Greenwald, J.W., Records, A.H., Nettleton, D., Lindow, S.E., Gross, D.C., and Beattie, G.A. (2013). Transcriptional responses of *Pseudomonas syringae* to growth in epiphytic versus apoplastic leaf sites. Proc. Natl. Acad. Sci. U. S. A. *110*, E425–34.

Zakataeva, N.P., Kutukova, E.A., Gronskiy, S. V., Troshin, P. V., Livshits, V.A., and Aleshin, V. V. (2006). Export of metabolites by the proteins of the DMT and RhtB families and its possible role in intercellular communication. Microbiology *75*, 438–448.

Zhang, J., Li, W., Xiang, T., Liu, Z., Laluk, K., Ding, X., Zou, Y., Gao, M., Zhang, X., Chen, S., et al. (2010a). Receptor-like Cytoplasmic Kinases Integrate Signaling from Multiple Plant Immune Receptors and Are Targeted by a *Pseudomonas syringae* Effector. Cell Host Microbe *7*, 290–301.

Zhang, W., Heemstra, J.R., Walsh, C.T., and Imker, H.J. (2010b). Activation of the pacidamycin PacL adenylation domain by MbtH-like proteins. Biochemistry *49*, 9946–9947.

Zhang, Z., Wu, Y., Gao, M., Zhang, J., Kong, Q., Liu, Y., Ba, H., Zhou, J., and Zhang, Y. (2012). Disruption of PAMP-induced MAP kinase cascade by a *Pseudomonas syringae* effector activates plant immunity mediated by the NB-LRR protein SUMM2. Cell Host Microbe *11*, 253–263.

Zhao, Y., Ma, Z., and Sundin, G.W. (2005). Comparative genomic analysis of the pPT23A plasmid family of *Pseudomonas syringae*. J. Bacteriol. *187*, 2113–2126.

Zhao, Y., Wu, J., Yang, J., Sun, S., Xiao, J., and Yu, J. (2012). PGAP: pan-genomes analysis pipeline. Bioinformatics *28*, 416–418.

Zhou, J.-M., and Chai, J. (2008). Plant pathogenic bacterial type III effectors subdue host responses. Curr. Opin. Microbiol. *11*, 179–185.

Appendix 1: High-Quality Draft Genome Sequence of *Pseudomonas syringae* Pathovar *syringae* Strain SM Isolated from Wheat

Alexey Dudnik[1] and Robert Dudler[1,2]

Genome Announcements

1(4):e00610-13

doi: 10.1128/genomeA.00610-13

Published: 15 August 2013

[1] Institute of Plant Biology, University of Zurich, Zollikerstrasse 107, CH-8008 Zurich, Switzerland

[2] Corresponding author

A1.1. Abstract

Pseudomonas syringae is one of the most wide-spread plant pathogens that can cause significant damage to crop plantations. Here, we announce a noncontiguous finished genome sequence of *Pseudomonas syringae* pv. *syringae* SM isolated from hexaploid wheat. The genome sequence revealed the smallest described complement of type III effectors.

A1.2. Genome announcement

Pseudomonas syringae strains have been isolated from >180 host species across the entire plant kingdom, including many agriculturally important crops (Bradbury, 1986)**.** The observed wide host range is reflected in a relatively large genetic heterogeneity among different pathovars. This is most pronounced in the complement of virulence factors, which is also assumed to be the key factor defining the host specificity (Lindeberg et al., 2009). *Pseudomonas syringae* pv. *syringae* strain SM was isolated from hexaploid wheat (*Triticum aestivum*) in the USA (Smith and Métraux, 1991). The strain, which was also denoted as D20, has been deployed in several studies addressing the issue of bacterium-induced systemic resistance in plants (Ramel et al., 2009; Rasmussen et al., 1991; Smith and Métraux, 1991; Smith et al., 1991; Vincent and Fulbright, 1983), but never as an infection model for wheat.

A 3-kb paired-end library was generated and sequenced at the Functional Genomics Center Zurich on a Roche Genome Sequencer FLX+ platform. A total of 974,051 quality filtered reads with a total of 213,333,037 bases were obtained, resulting in 34.8-fold average sequencing coverage. The obtained reads were further *de novo* assembled using Newbler 2.5.3 into 64 contigs combined into one 6.08 Mb-long super-scaffold and 3 smaller scaffolds (46.5 kb, 9.09 kb and 5.24 kb in size). The largest of the minor scaffolds turned out to be a pPT23A family plasmid, the 9-kb scaffold showed sequence similarity to non-ribosomal peptide synthase (NRPS) modules, while the smallest scaffold constituted a rRNA operon. A portion of intra-scaffold gaps was closed by sequencing of PCR products using Sanger technology, decreasing the total number of contigs to 26. However, it was not possible to precisely map the 9-kb scaffold, but due to insignificance to the project it was excluded from the assembly. Initial open-reading frame (ORF) prediction and functional annotation has been performed using the RAST server (Aziz et al., 2008). The start codons of all the predicted ORFs were further manually verified using the position of potential ribosomal binding sites and BLASTP (Altschul et al., 1990) alignments with homologous ORFs from other *Pseudomonas* strains as a

reference. Functional annotations have as well been refined for every ORF using BLASTP searches against the nonredundant protein sequence database (nr) and the NCBI Conserved-Domain search engine (Marchler-Bauer et al., 2011).

The estimated genome size of strain SM is 6,124,102 base pairs with an average G+C content of 58.73%. It encodes 5,072 protein-coding sequences (excluding pseudogenes), five rRNA operons, and 64 tRNA genes for all of the amino acids. Notably, it contains a complete type III secretion system and seven known effector proteins: AvrE1, HopAA1, HopI1, HopM1, HopBA1, HopA2, and HopAZ1. In addition, there are two complete type VI secretion system gene clusters and 12 putative effector proteins belonging to the VgrG and Hcp1 families, as well as intact gene clusters for biosynthesis of syringopeptin and mangotoxin. All of these genome components have previously been demonstrated to be involved in virulence and epiphytic fitness of *P. syringae*, as well as in competition of pseudomonads with other microbial species (Arrebola et al., 2009; Bender et al., 1999; Cunnac et al., 2011; Haapalainen et al., 2012; Hood et al., 2010; Lindeberg et al., 2012).

A1.3. Nucleotide sequence accession numbers

This Whole Genome Shotgun project has been deposited at DDBJ/EMBL/GenBank under the accession APWT00000000. The version described in this paper is the first version, APWT01000000. The assigned NCBI taxonomy identification number is 1262350.

Appendix 2: Noncontiguous-Finished Genome Sequence of *Pseudomonas syringae* Pathovar *syringae* Strain B64 Isolated from Wheat

Alexey Dudnik[1] and Robert Dudler[1,2]

Standards in Genomic Sciences

8(3):420-429

doi: 10.4056/sigs.3997732

Published: 30 July 2013

[1] Institute of Plant Biology, University of Zurich, Zollikerstrasse 107, CH-8008 Zurich, Switzerland

[2] Corresponding author

A2.1. Abstract

The Gram-negative gammaproteobacterium *Pseudomonas syringae* is one of the most wide-spread plant pathogens and has been repeatedly reported to cause significant damage to crop plantations. Research on this pathogen is very intensive, but most of it is done on isolates that are pathogenic to *Arabidopsis*, tomato, and bean. Here, we announce a high-quality draft genome sequence of *Pseudomonas syringae* pv. *syringae* B64 which is the first published genome of a *P. syringae* strain isolated from wheat up to date. The genome sequence will assist in gaining insights into basic virulence mechanisms of this pathogen which has a relatively small complement of type III effectors.

A2.2. Introduction

Pseudomonas syringae strains have been isolated from more than 180 host species (Bradbury, 1986) across the entire plant kingdom, including many agriculturally important crops, such as bean, tomato, cucumber, as well as kiwi, stone fruit, and olive trees. Strains are divided into more than 50 pathovars primarily based on host-specificity, disease symptoms, and biochemical profiles (González et al., 2000; Qi et al., 2011; Young, 2010). The first strain of this species was isolated from a lilac tree (*Syringa vulgaris*), which gave origin to its name (Palleroni, 2005a).The observed wide host range is reflected in a relatively large genetic heterogeneity among different pathovars. This is most pronounced in the complement of virulence factors, which is also assumed to be the key factor defining host specificity (Lindeberg et al., 2009). For successful survival and reproduction, both epiphytic and endophytic *P. syringae* strains deploy different sets of type III and type VI secretion system effectors, phytotoxins, EPS, and other types of secreted molecules (Bender et al., 1999; Denny, 1995; Haapalainen et al., 2012; Lindeberg et al., 2009, 2012; Schellenberg et al., 2010). Currently, there are three completely sequenced *P. syringae* genomes published: pathovar *syringae* strain B728a which causes brown spot disease of bean (Feil et al., 2005), pathovar *tomato* strain DC3000 which is pathogenic to tomato and *Arabidopsis* (Buell et al., 2003), and pathovar *phaseolicola* strain 1448A, causal agent of halo blight on bean (Joardar et al., 2005). There are also a number of incomplete genomes of various qualities available for other strains.

Pseudomonas syringae pv. *syringae* strain B64 was isolated from hexaploid wheat (*Triticum aestivum*) in Minnesota, USA (Denny, 1988). The strain has been deployed in several studies mainly addressing phylogenetic diversity of *P. syringae* varieties (Denny, 1988; Denny et al., 1988; Hwang et al., 2005; Sundin et al., 1994), but never as an infection model for wheat. The genome sequencing of the B64 strain and its comparison with the other published genomes should reveal wheat-specific adaptations and give insights in virulence strategies for colonizing monocot plants.

A2.3. Classification and features

Pseudomonas syringae belongs to class *Gammaproteobacteria*. Detailed classification of this species is still under heavy debate. Young and colleagues have proposed to group all plant-pathogenic oxidase-negative and fluorescent *Pseudomonas* strains into a single species, *P. syringae*, which is to be further sub-divided into pathovars (Young, 2010; Young et al., 1978). Several DNA hybridization studies have shown a large genetic heterogeneity among the groups, however biochemical characteristics, with a few exceptions, did not allow elevating those into distinct species (Gardan et al., 1992, 1999). Currently, the species is divided into five phylogenetic clades based on MLST analysis. *P. syringae* pv. *syringae* (*Psy*) strains belong to group II within this nomenclature (Baltrus et al., 2011). The basic characteristics of *Psy* B64 are summarized in Table 1, while its phylogenetic position is depicted in Figure 1.

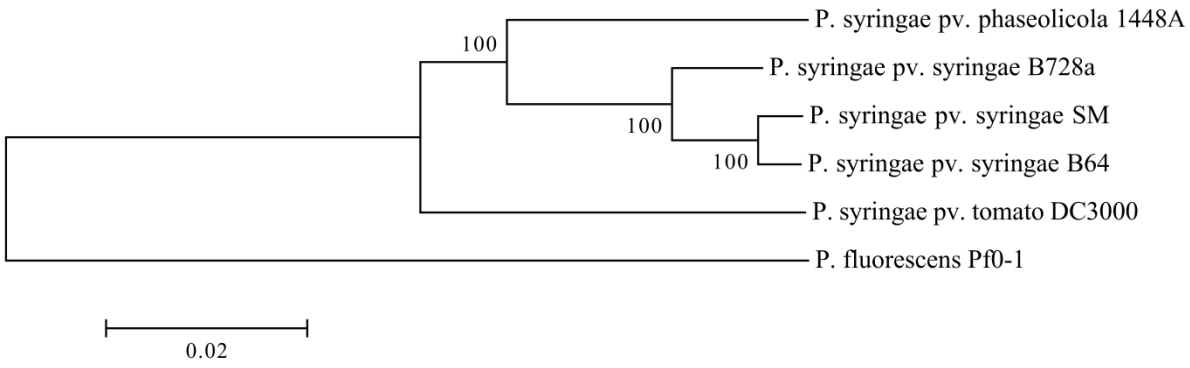


Figure 1. Phylogenetic tree constructed using neighbor-joining method using MLST approach (Sarkar and Guttman, 2004) and MEGA 5.10 software suit (Tamura et al., 2011) with 1,000 bootstraps. The tree features the three completely sequenced *P. syringae* model strains *Pto* DC3000, *Psy* B728a, and *Pph* 1448A, the strain *Psy* B64 itself, as well as another wheat-isolated strain *Psy* SM. The model strains represent the three major phylogenetic clades of *P. syringae*: I, II and III respectively. *P. fluorescens* Pf0-1 was used as an outgroup. The analysis confirms placement of *Psy* B64 into clade II.

Table 1. Classification and the general features of *P. syringae* pv. *syringae* B64 according to the MIGS recommendations (Field et al., 2008)

MIGS ID	Property	Term	Evidence code[a]
		Domain *Bacteria*	TAS
		Phylum *Proteobacteria*	TAS
		Class *Gammaproteobacteria*	TAS
	Current classification	Order *Pseudomonadales*	TAS
		Family *Pseudomonadaceae*	TAS
		Genus *Pseudomonas*	TAS
		Species *Pseudomonas syringae*	TAS
		Pathovar *syringae*	TAS
		Strain B64	TAS
	Gram stain	Negative	TAS
	Cell shape	Rod-shaped	TAS
	Motility	Motile	TAS
	Sporulation	None	TAS
	Temperature range	Mesophilic	TAS
	Optimum temperature	28°C	TAS
MIGS-22	Oxygen	Aerobic	TAS
	Carbon source	Heterotrophic	TAS
	Energy metabolism	Chemoorganotrophic	TAS
MIGS-6	Habitat	Host-associated	TAS
MIGS-6.3	Salinity	Not reported	
MIGS-10	Extrachromosomal elements	None	IDA
MIGS-11	Estimated Size	5.93 Mb	IDA
MIGS-15	Biotic relationship	Parasitic	NAS
MIGS-14	Pathogenicity	Pathogenic	TAS
	Host	*Triticum aestivum*	TAS
	Host taxa ID	4565	
	Cell arrangement	Single	TAS
	Biosafety level	1	NAS
	Isolation source	Leaf	NAS
MIGS-4	Geographic location	Minnesota, USA	TAS

[a]Evidence codes - IDA: Inferred from Direct Assay (first time in publication); TAS: Traceable Author Statement (Denny, 1988; Garrity et al., 2005; Hwang et al., 2005; Palleroni, 2005b; Sundin et al., 1994; Woese, 1990; Young et al., 1977, 1978); NAS: Non-traceable Author Statement (not directly observed for the living, isolated sample, but based on a generally accepted property of the species, or anecdotal evidence). These evidence codes are from the Gene Ontology project (Ashburner et al., 2000).

Psy B64 has similar physiological properties as other representatives of its genus. It can grow in complex media such as LB (Bertani, 1951) or King's B (King et al., 1954), as well as in various defined minimal media: HSC (Palmer and Bender, 1993), MG-agar (Keane et al.), PMS (Gasson, 1980), AB-agar (Clark and Maaløe, 1967), and SRM_{AF} (Mo and Gross, 1991). Even though the optimal growth temperature is 28°C, the bacterium can also replicate at 4°C. Growth is completely inhibited above 35°C. *Psy* B64 is capable of endophytic growth in the wheat leaf mesophyll, but does not seem to cause any symptoms unless a very high inoculation dose is applied.

The bacterium has a weak resistance to ampicillin (25 mg/L) and chloramphenicol (10 mg/L). It is also possible to develop spontaneous rifampicin-resistant mutants. In addition, the genomic sequence predicts this strain to be polymyxin B insensitive due to presence of the *arn* gene cluster.

A2.4. Genome sequencing information

A2.4.1. Genome project history

The organism was selected for sequencing because it has been identified to have a syringolin biosynthesis gene cluster (Amrein et al., 2004). Syringolin is a proteasome inhibitor produced by some strains of pathovar *syringae*. As a consequence of proteasome inactivation a number of plant intracellular pathways are being inhibited, including the entire salicylic acid-dependent defense pathway, thus promoting the entry of bacteria into leaf tissue and subsequent endophytic growth (Schellenberg et al., 2010). Since up to now, it has not been possible to establish an infection model for syringolin in the model plant *Arabidopsis*, it was decided to explore another common research target and one of the most important crop plants, bread wheat (*Triticum aestivum*). The genome project has been deposited in the Genbank Database (ID 180994) and the genome sequence is available under accession number ANZF00000000. The version described in this paper is the first version, ANZF01000000. The details of the project are shown in Table 2.

A2.4.2. Growth conditions and DNA isolation

P. syringae pv. *syringae* strain B64 was grown in 40 mL of LB medium at 28°C, 220 rpm until OD_{600} of ~1.0. Genomic DNA was isolated from the pelleted cell using a Qiagen Genomic-tip 100/G column (Qiagen, Hilden, Germany) according to the manufacturer's instructions.

Table 2. Genome sequencing project information

MIGS ID	Property	Term
MIGS-31	Finishing quality	High-quality draft
MIGS-28	Libraries used	3kb paired-end library
MIGS-29	Sequencing platform	Roche Genome Sequencer FLX+
MIGS-31.2	Sequencing coverage	31.8×
MIGS-30	Assembler	Newbler 2.5.3
MIGS-31.3	Contigs	41
MIGS-32	Gene calling method	RAST server
	NCBI project ID	180994
	NCBI accession number	ANZF00000000
	Date of release	January 18, 2013
	GOLD ID	Gc02493
	Database: IMG/ER	2523533564
	Project relevance	Plant-pathogen interactions, infection model for syringolin

A2.4.3. Genome sequencing and assembly

A 3kb paired-end library was generated and sequenced at the Functional Genomics Center Zurich on a Roche Genome Sequencer FLX+ platform. A total of 872,570 high-quality filtered reads with a total of 188,465,376 bases were obtained, resulting in 31.8-fold average sequencing coverage. The obtained reads were assembled *de novo* using Newbler 2.5.3. This resulted in 150 contigs combined into one 6 Mb-long super-scaffold and 3 smaller scaffolds of 5.29 kb, 2.84 kb and 2.74 kb in size. The largest of the minor scaffolds constituted a ribosomal RNA operon, the other two showed sequence similarity to non-ribosomal peptide synthase modules. A portion of intra-scaffold gaps have been closed by sequencing of PCR products using Sanger technology, decreasing the total number of contigs to 41 with a contig N50 value of 329.4 kb, the longest contig being 766.5 kb long. Note that the Genbank record contains 42 contigs due to fact that one of the contigs was split into two parts in order to start the assembly with the *dnaA* gene. While closing gaps it became possible to allocate the positions of all ribosomal operons by sequence overlap and thus to incorporate the largest of the minor scaffolds. However, it was not possible to precisely map the remaining two minor scaffolds. These must be located within two distinct remaining large gaps, but due to insignificance to the project they have been excluded from the assembly.

A2.4.4. Genome annotation

Initial open-reading frame (ORF), tRNA, and rRNA prediction and functional annotation has been performed using the RAST (Rapid Annotation using Subsystem Technology) server (Aziz et al., 2008). For the purpose of comparison, the genome has also been annotated using Prokka (Seemann, 2013), which utilizes Prodigal (Hyatt et al., 2010) for ORF prediction (the RAST server utilizes a modified version of Glimmer (Delcher et al., 2007)). Start codons of all the predicted ORFs were further verified manually, using the position of potential ribosomal binding sites and BLASTP (Altschul et al., 1990) alignments with homologous ORFs from other *P. syringae* strains as a reference. Functional annotations have also been refined for every ORF using BLASTP searches against the non-redundant protein sequence database (nr) and the NCBI Conserved-Domain search engine (Marchler-Bauer et al., 2011). Functional category assignment and signal peptide prediction was done using the Integrated Microbial Genomes/Expert reviews (IMG/ER) system (Markowitz et al., 2009).

A2.4.5. Genome properties

The genome of the strain B64 is estimated to be comprised of 5,930,035 base pairs with an average GC-content of 58.55 % (Table 3 and Figure 2), which is similar to what is observed in other *P. syringae* strains (Buell et al., 2003; Feil et al., 2005; Joardar et al., 2005). Of the 5,021 predicted genes, 4,947 were protein coding genes, 4 ribosomal RNA operons, and 61 tRNA genes; 78 were identified to be pseudo-genes. The majority of the protein-coding genes (83.65 %) were assigned a putative function, while the remaining ones were annotated as hypothetical proteins. The distribution of genes into COGs functional categories is presented in Table 4.

The genome contains a complete canonical type III secretion system and ten known effector proteins: AvrE1, HopAA1, HopAG1, HopAH1, HopAI1, HopAZ1, HopBA1, HopI1, HopM1, and HopZ3. Out of these ten, the first five are present in all other sequenced *P. syringae* strains, thereby constituting the effector core, whereas the latter five could be host-determinants for wheat. That there is such a small number of effectors is not something unusual, and is seen in other strains of clade II (Baltrus et al., 2011). In addition, there are two complete type VI secretion system gene clusters and nine putative effector proteins belonging to the VgrG and Hcp1 families. *Psy* B64 genome also encodes gene clusters for biosynthesis of four phytotoxin: syringomycin, syringopeptin, syringolin, and mangotoxin. All of the above-mentioned genome components have

been previously demonstrated to be involved in virulence, epiphytic fitness of *P. syringae*, as well as in competition with other microbial species (Arrebola et al., 2009; Bender et al., 1999; Cunnac et al., 2011; Groll et al., 2008; Haapalainen et al., 2012; Lindeberg et al., 2012; Schellenberg et al., 2010). Additional identified virulence-associated traits are: exopolysaccharides alginate, Psl, and levan biosynthesis, surfactant syringofactin, type VI pili, large surface adhesins, siderophores pyoverdine and achromobactin, proteases and other secreted hydrolytic enzymes, RND-type transporters (including putative *mexAB*, *mexCD*, *mexEF*, and *mexMN* homologs (Mima et al., 2005; Poole, 2001)), all of which are found in other *P. syringae* strains. It is also notable that *inaZ* gene encoding ice-nucleation protein is truncated by a frameshift, thus making this strain ice-negative. The latter contradicts results of a previous study by Hwang and colleagues (Hwang et al., 2005) in which *Psy* B64 has been identified to be ice-positive. This could be due to an assembly error, or the frameshift could have been introduced at a later point during propagation.

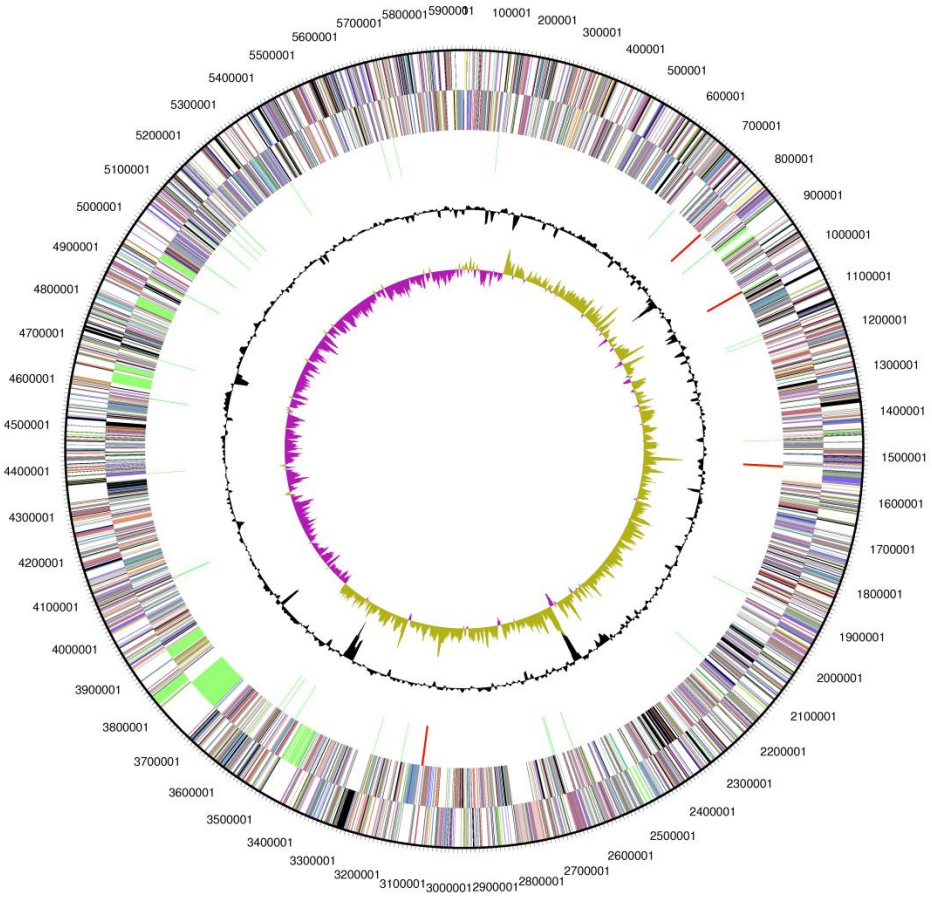


Figure 2. Graphical map of the chromosome. From outside to the center: genes on forward strand (colored by COG categories), genes on reverse strand (colored by COG categories), RNA genes: tRNAs - green, rRNAs - red, other RNAs - black, GC content, and GC skew

Table 3. Genome Statistics

Attribute	Value	% of Total
Estimated genome size (bp)	5,930,035	100.00
Estimated total gap length (bp)	56,737	0.96
DNA coding region (bp)	5,146,184	86.78
DNA G+C content (bp)	3,472,195	58.55
Number of replicons	1	-
Extra-chromosomal elements	0	-
Total genes	5,021	100.00
Protein-coding genes[a]	4,869	96.97
RNA genes	74	1.47
rRNA genes	13	0.26
5S rRNA	5	0.10
16S rRNA	4	0.08
23S rRNA	4	0.08
tRNA genes	61	1.21
rRNA operons	4	-
Pseudo-genes	78	1.55
Protein coding genes with function prediction	4,200	83.65
without function prediction[a]	669	13.32
Protein coding genes with COGs	4,013	79.92
with KOGs	1,698	33.82
with Pfam	4,256	84.76
with TIGRfam	1,641	32.68
in paralog clusters	3,933	78.33
Proteins with signal peptides	511	10.18
Proteins with transmembrane helices	1,112	22.15

[a]excluding pseudo-genes; therefore percentage values differ from those displayed at the IMG/ER

Table 4. Number of genes associated with the 25 general COG functional categories

Code	Value	%	Description
J	208	4.60	Translation
A	1	0.02	RNA processing and modification
K	352	7.78	Transcription
L	162	3.58	Replication, recombination and repair
B	1	0.02	Chromatin structure and dynamics
D	42	0.93	Cell cycle control, mitosis and meiosis
Y	-	-	Nuclear structure
V	48	1.06	Defense mechanisms
T	332	7.34	Signal transduction mechanisms
M	275	6.08	Cell wall/membrane biogenesis
N	159	3.52	Cell motility
Z	1	0.02	Cytoskeleton
W	-	-	Extracellular structures
U	140	3.10	Intracellular trafficking and secretion
O	156	3.45	Posttranslational modification, protein turnover, chaperones
C	230	5.09	Energy production and conversion
G	265	5.86	Carbohydrate transport and metabolism
E	439	9.71	Amino acid transport and metabolism
F	90	1.99	Nucleotide transport and metabolism
H	177	3.91	Coenzyme transport and metabolism
I	152	3.36	Lipid transport and metabolism
P	268	5.93	Inorganic ion transport and metabolism
Q	122	2.70	Secondary metabolites biosynthesis, transport and catabolism
R	515	11.39	General function prediction only
S	388	8.58	Function unknown
-	1,008	20.08	Not in COGs